



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 111121

TO: Minh-Tam Davis
Location: cm1/8a01/8e12
Art Unit: 1642
Friday, January 02, 2004

Case Serial Number: 09/554945

From: Mary Jane Ruhl
Location: Biotech-Chem Library
CM1-6A06
Phone: 605-1155

maryjane.ruhl@uspto.gov

Search Notes

Examiner Davis,

Here are the results for your recent search request.

Please feel free to contact me if you have any questions about these results.

Thank you for using STIC services. We appreciate the opportunity to serve you.

Sincerely,

Mary Jane Ruhl
Technical Information Specialist
STIC
CM-1, Rm. 6-A-06
605-1155

THIS PAGE BLANK (USPTO)

PN WO200140466-A2.
 XX 07-JUN-2001.
 XX 01-DEC-2000; 2000WO-US32678.
 XX 01-DEC-1999; 99WO-US28301.
 XX 01-DEC-1999; 99WO-US28634.
 XX 02-DEC-1999; 99WO-US28551.
 XX 02-DEC-1999; 99WO-US28564.
 XX 02-DEC-1999; 99WO-US28565.
 XX 02-DEC-1999; 99US-0170262.
 XX 16-DEC-1999; 99WO-US30095.
 XX 20-DEC-1999; 99WO-US30911.
 XX 20-DEC-1999; 99WO-US30999.
 XX 06-JAN-1999; 99WO-US31243.
 XX 06-JAN-2000; 2000WO-US00277.
 XX 06-JAN-2000; 2000WO-US00376.
 XX 11-FEB-2000; 2000WO-US03565.
 XX 18-FEB-2000; 2000WO-US04341.
 XX 18-FEB-2000; 2000WO-US04342.
 XX 22-FEB-2000; 2000WO-US04414.
 XX 24-FEB-2000; 2000WO-US04914.
 XX 24-FEB-2000; 2000WO-US05004.
 XX 01-MAR-2000; 2000WO-US05601.
 XX 20-MAR-2000; 2000WO-US07377.
 XX 21-MAR-2000; 2000WO-US07532.
 XX 30-MAR-2000; 2000WO-US08439.
 XX 17-MAY-2000; 2000WO-US13705.
 XX 22-MAY-2000; 2000WO-US14042.
 XX 30-MAY-2000; 2000WO-US14941.
 XX 02-JUN-2000; 2000WO-US15264.
 XX 10-NOV-2000; 2000WO-US30873.
 XX (GETH) GENENTECH INC.
 XX Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX WPI; 2001-408281/43.
 XX N-PSDB; AAS21288.
 XX Isolated, secretory and transmembrane PRO polypeptide used to detect
 PT other PRO polypeptides, link bioactive molecules to cells expressing
 PT PRO polypeptides, and detect the presence of mammalian tumours e.g.
 PT lung, breast, prostate, cervical -
 XX Claim 12; Fig 90; 813pp; English.
 XX AAU12172-AAU12446 represent novel human secretory and transmembrane
 CC PRO polypeptides. The PRO polypeptides are useful to detect other
 CC PRO polypeptides, to link bioactive molecules to cells expressing
 CC PRO polypeptides, to modulate biological activities of cells expressing
 CC PRO polypeptides, and to detect the presence of mammalian lung, colon,
 CC breast, prostate, rectal, cervical or liver tumours by comparing PRO
 CC polypeptide expression in a cell sample to that in a control sample.
 CC Some of the 275 sequences are also useful to stimulate the release of
 CC tumour necrosis factor-alpha (TNF-alpha) from human blood, the
 CC proliferation or differentiation of chondrocytes, the proliferation or
 CC gene expression in pericyte cells, the release of proteoglycans from
 CC cartilage, the proliferation of inner ear utricular supporting cells or
 CC of T-lymphocytes, the release of a cytokine from peripheral blood
 CC monocytes (PMNCs), or the proliferation of endothelial cells. Some of
 CC the PRO polypeptides may modulate glucose or free fatty acid uptake by
 CC skeletal muscle cells or by adipocytes; or inhibit binding of A-peptide
 CC to factor VIIA. The PRO polypeptides can be used in assays to identify
 CC molecules involved in binding interactions. The polynucleotides encoding
 CC PRO polypeptides can be used to generate probes, antisense RNA/DNA,
 CC transgenic or knock out animals and can be used in gene therapy.
 XX Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 22; Length 468;
 Best Local Similarity 100.0%; Pred. No. 4.1e-146;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGFLGTGTWILVLVLPQAPKPGSQDQKSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
 DB 1 MGFLGTGTWILVLVLPQAPKPGSQDQKSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
 QY 61 ENKPGQSNYSFVDNLNLLKAITKEKTEKERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
 DB 61 ENKPGQSNYSFVDNLNLLKAITKEKTEKERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
 QY 121 DSTKSGLDHKKFQDDPGDLHQLDGTPLTAEDIVHKAARIYEENDRAVFDKIVSKLLNLGL 180
 DB 121 DSTKSGLDHKKFQDDPGDLHQLDGTPLTAEDIVHKAARIYEENDRAVFDKIVSKLLNLGL 180
 QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKIPEKVTMPMAAIODGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKIPEKVTMPMAAIODGL 240
 QY 241 AKGENDETQSVNTLTTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKAEKETLI 300
 DB 241 AKGENDETQSVNTLTTLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKAEKETLI 300
 QY 301 TIMKTILDFVMMVKYGTISPEBGVSYLENDEMIALQTKNKLKKNATDNISKLPAPSE 360
 DB 301 TIMKTILDFVMMVKYGTISPEBGVSYLENDEMIALQTKNKLKKNATDNISKLPAPSE 360
 QY 361 KSHEETDSTKEEAAKMEKYGSLKSDTKDSDNSPGGKTDRPKGTEAYLEAIRKNIEWLK 420
 DB 361 KSHEETDSTKEEAAKMEKYGSLKSDTKDSDNSPGGKTDRPKGTEAYLEAIRKNIEWLK 420
 QY 421 KHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
 DB 421 KHDKKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
 RESULT 3
 AAB87600
 ID AAB87600 standard; Protein; 468 AA.
 XX AAB87600;
 XX 15-MAY-2001 (first entry)
 XX Human PRO5990.
 XX Human; PRO protein; mapping.
 XX Homo sapiens.
 XX WO200116318-A2.
 XX 08-MAR-2001.
 XX 24-AUG-2000; 2000WO-US23328.
 XX 01-SEP-1999; 99WO-US20111.
 XX 15-SEP-1999; 99WO-US21090.
 XX 07-DEC-1999; 99US-0169495.
 XX 09-DEC-1999; 99US-0170262.
 XX 11-JAN-2000; 2000US-0175481.
 XX 18-FEB-2000; 2000WO-US04341.
 XX 18-FEB-2000; 2000WO-US04342.
 XX 22-FEB-2000; 2000WO-US04414.
 XX 01-MAR-2000; 2000WO-US05601.
 XX 03-MAR-2000; 2000US-0187202.
 XX 25-APR-2000; 2000US-0199397.
 XX 22-MAY-2000; 2000WO-US14042.
 XX 05-JUN-2000; 2000US-0209832.
 XX (GETH) GENENTECH INC.
 XX PA
 XX

PI	Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI	Grimaldi CU, Gurney AL, Watanabe CK, Wood WI;
XX	WPI; 2001-183260/18.
DR	N-PSDB; AAF92132.
XX	
PT	Eighty four nucleic acids encoding PRO polypeptides, useful in
PT	molecular biology, including use as hybridization probes, and in
PT	chromosome and gene mapping. -
XX	
PS	Claim 12; Fig 150; 278pp; English.
XX	
CC	The present sequence is a human PRO polypeptide (secreted and
CC	transmembrane). The PRO protein, and PRO agonists, PRO antagonists or
CC	anti-PRO antibodies are useful for preparation of a medicament useful in
CC	the treatment of a condition which is responsive to the PRO protein,
CC	agonists, antagonists or anti-PRO antibodies. The PRO protein may also be
CC	employed as molecular weight markers for protein electrophoresis. The PRO
CC	coding sequence has applications in molecular biology, including use as
CC	hybridisation probes, and in chromosome and gene mapping.
XX	
SQ	Sequence 468 AA;
	Query Match 100.0%; Score 2399; DB 22; Length 468;
	Best Local Similarity 100.0%; Pred. No. 4.1e-146;
	Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MGFLGTGTLVLVLPIQAFPPGGSQDKSLHNRELSAERPLNQLIAEAEEDKIKKTYPP 60
DB	
QY	1 MGFLGTGTLVLVLPIQAFPPGGSQDKSLHNRELSAERPLNQLIAEAEEDKIKKTYPP 60
DB	
QY	61 ENKPGQSNYSFVDNLNLLKATTEKEKTEKERSQSRSSPLDNKLNVEDVSTNKRKLIDDY 120
DB	
QY	61 ENKPGQSNYSFVDNLNLLKATTEKEKTEKERSQSRSSPLDNKLNVEDVSTNKRKLIDDY 120
DB	
QY	121 DSTKSGLDHFKFQDDPDGHLQDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLNLGL 180
DB	
QY	121 DSTKSGLDHFKFQDDPDGHLQDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLNLGL 180
DB	
QY	181 ITBSQAHTLDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKIEKVTMPAAIIOGL 240
DB	
QY	181 ITBSQAHTLDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKIEKVTMPAAIIOGL 240
DB	
QY	241 AKGENDETVENTLTLTNGLERRTKTSYSEDNFEELQYFPNFYALLKSIDSEKAKEKETLI 300
DB	
QY	241 AKGENDETVENTLTLTNGLERRTKTSYSEDNFEELQYFPNFYALLKSIDSEKAKEKETLI 300
DB	
QY	301 TIMKTLIDFVKVMVKYGTISPEGVSVLENLDEMIALQTNKLEKKNATDNISKLPAPSE 360
DB	
QY	301 TIMKTLIDFVKVMVKYGTISPEGVSVLENLDEMIALQTNKLEKKNATDNISKLPAPSE 360
DB	
QY	361 KSHEETDSTKEEAANKMEKEYGSLKSDTKDNSNPGGKTDEFPKGKTEAYLEAIRKNIEWLK 420
DB	
QY	361 KSHEETDSTKEEAANKMEKEYGSLKSDTKDNSNPGGKTDEFPKGKTEAYLEAIRKNIEWLK 420
DB	
QY	421 KHDKGNKEDYDLKMRDFTINKQADAYVEKGIIDKEEAIAIKRIYSSL 468
DB	
QY	421 KHDKGNKEDYDLKMRDFTINKQADAYVEKGIIDKEEAIAIKRIYSSL 468
DB	
RESULT 4	
ID	ABG95925
AC	ABG95925 standard; Protein; 468 AA.
XX	
AC	ABG95925;
XX	
DT	10-DEC-2002 (first entry)
XX	
XX	Human secreted/transmembrane protein PRO5990.
XX	
KW	Human; secreted protein; transmembrane protein; antirheumatic;
KW	antiarthritic; osteopathic; sports-related joint problem;
KW	articular cartilage defect; osteoarthritis; rheumatoid arthritis.

CC polypeptide, H is a PRO20233 polypeptide and I is a PRO1890
CC polypeptide. The sample comprises a cell suspected of expressing the A,
CC B, C or D polypeptide. The E, F, G, H or I polypeptide is labeled with
CC a detectable label or is attached to a solid support. The proteins are
CC useful for linking a bioactive molecule to a cell expressing a
CC polypeptide designated as A, B, C or D or E, F, G, H or I. The bioactive
CC molecule is a toxin, a radiolabel or an antibody. The bioactive molecule
CC causes death of the cell. A, B, C, D, E, F, G, H, or I, or antibodies
CC against them are useful for modulating a biological activity of a cell
CC expressing a polypeptide designated as A, B, C or D or E, F, G, H, or
CC I. The cell is killed. The proteins are useful for identifying
CC agonists or antagonists, for the preparation of a medicament useful in
CC the treatment of a condition which is responsive to the proteins, as
CC molecular weight markers for protein electrophoresis purposes, and as
CC therapeutic agents for treating sports-related joint problems,
CC articular cartilage defects, osteoarthritis or rheumatoid arthritis.
CC Nucleic acids encoding the proteins are useful as hybridisation probes,
CC in chromosome and gene mapping, in the generation of anti-sense RNA and
CC DNA, for the preparation of the proteins, to generate transgenic or
CC knockout animals which are useful in the development and screening of
CC therapeutic useful reagents, for chromosome identification, and in gene
CC therapy. The antibody is useful as a therapeutic agent, in a diagnostic
CC assay and for affinity purification of the protein from recombinant
CC cell culture natural sources. The present sequence represents a novel
CC secreted or transmembrane protein of the invention.

Query Match 100.0%; Score 2399; DB 23; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGFLGTGFWILVLPQAPFKPGSQDKSLHNSLSAERPLNQIAEAEEDKIKTYPP 60
Db 1 MGFLGTGFWILVLPQAPFKPGSQDKSLHNSLSAERPLNQIAEAEEDKIKTYPP 60
Qy 61 ENKPGQSNYSFVNLNLLKATTEKEKEKESQSRSSPLDNKLNVEDVDSTKRNKLIDDY 120
Db 61 ENKPGQSNYSFVNLNLLKATTEKEKEKESQSRSSPLDNKLNVEDVDSTKRNKLIDDY 120
Qy 121 DSTKSGLDHKEFDQDPDGLHLDGTPLTAEIVHKIARIYEENDRAVDFDKIVSKLLNLGL 180
Db 121 DSTKSGLDHKEFDQDPDGLHLDGTPLTAEIVHKIARIYEENDRAVDFDKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPEKVTIPMAAIQDGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPEKVTIPMAAIQDGL 240
Qy 241 AKGENDETIVSNTLITNGLERTKTYSDNPFEELOYPNFPVALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETIVSNTLITNGLERTKTYSDNPFEELOYPNFPVALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLIDFVQMWKYGTISPPEGVSYLENDEMIALQTNKLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVQMWKYGTISPPEGVSYLENDEMIALQTNKLEKNATDNISKLPAPSE 360
Qy 361 KSHETDSTKEEAARKEKEYGSLDSTKDDNSNPGKTDPEPKGTETAYLEAIRKNIEWLK 420
Db 361 KSHETDSTKEEAARKEKEYGSLDSTKDDNSNPGKTDPEPKGTETAYLEAIRKNIEWLK 420
Qy 421 KHKKKGNKEDYDLKMRDFINKQADAVVEKGIILDKERAEAIKRIYSSL 468
Db 421 KHKKKGNKEDYDLKMRDFINKQADAVVEKGIILDKERAEAIKRIYSSL 468

RESULT 5
ABU69097
ID ABU69097 standard; Protein; 468 AA.
XX AC ABU69097;
XX XX
DT 02-JUN-2003 (first entry)
XX

DE Human PRO polypeptide #18.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein; cancer;
KW non-insulin dependent diabetes mellitus; septic shock; stroke;
KW rheumatoid arthritis; graft-versus-host disease; cardiac ischaemia;
KW psoriasis; inflammatory bowel disease; asthma; anti-diabetic;
KW cytostatic; immunosuppressive; antirheumatic; antiarthritic;
KW cerebroprotective; vasotropic; antipsoriatic; antiinflammatory;
KW antiasthmatic.
OS Homo sapiens.
XX
PN US2003008348-A1.
XX
PD 09-JAN-2003.
XX
PF 26-DEC-2001; 2001US-0035855.
XX
PR 14-MAY-1999; 99WO-US10733.
PR 02-DEC-1999; 99WO-US28551.
PR 22-DEC-1999; 99WO-US30720.
PR 01-MAR-2000; 2000WO-US05601.
PR 02-MAR-2000; 2000WO-US05841.
PR 22-MAY-2000; 2000WO-US14042.
PR 02-JUN-2000; 2000WO-US15264.
PR 23-AUG-2000; 2000WO-US23522.
PR 24-AUG-2000; 2000WO-US23328.
PR 01-DEC-2000; 2000WO-US32678.
PR 01-JUN-2001; 2001WO-US17800.
PR 20-JUN-2001; 2001WO-US19892.
PR 29-JUN-2001; 2001WO-US21066.
PR 09-JUL-2001; 2001WO-US21735.
PR 15-MAY-1998; 98US-085579P.
PR 15-DEC-1998; 98US-112514P.
PR 22-DEC-1998; 98US-113300P.
PR 23-DEC-1998; 98US-113430P.
PR 23-DEC-1998; 98US-113605P.
PR 22-JAN-1999; 99US-116843P.
PR 23-MAR-1999; 99US-125774P.
PR 23-MAR-1999; 99US-125778P.
PR 24-MAR-1999; 99US-125826P.
PR 31-MAR-1999; 99US-127035P.
PR 05-APR-1999; 99US-127106P.
PR 13-APR-1999; 99US-129122P.
PR 21-APR-1999; 99US-130359P.
PR 25-MAY-1999; 99US-135750P.
PR 08-JUN-1999; 99US-138166P.
PR 20-JUL-1999; 99US-144791P.
PR 03-AUG-1999; 99US-146970P.
PR 29-OCT-1999; 99US-162506P.
PR 16-AUG-2001; 2001US-0931836.
XX
(GETH) GENENTECH INC.
XX
PI Desnoyers L, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
PI Stewart TA, Watanabe CK, Wood WI, Zhang Z;
XX
WPI: 2003-341326/32.
XX
N-PSDB; ACA06116.
XX
New PRO polypeptides and nucleic acid molecules, useful for diagnosing
PT or treating diabetes mellitus, cancers, septic shock, inflammatory
PT bowel disease or asthma, or in gene therapy, chromosome identification
PT or tissue typing
XX
PS Claim 12; Fig 36; 196pp; English.
XX
CC The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The
CC PRO polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful in diagnosing or
CC treating non-insulin dependent diabetes mellitus, cancers, septic
CC shock, rheumatoid arthritis, graft-versus-host disease, stroke, cardiac

```

CC ischaemia, psoriasis, inflammatory bowel disease or asthma. The PRO
CC polynucleotide sequences may be used as hybridisation probes in
CC chromosome and gene mapping, or in generating antisense RNA and DNA.
CC They are also useful in preparing PRO polypeptides, in assays to
CC identify other proteins or molecules involved in binding reaction, to
CC generate transgenic animals or knockout animals, which in turn are
CC useful in the development and screening of therapeutically useful
CC reagents, for chromosome identification, and tissue typing. The PRO
CC polypeptides and nucleic acid molecules are also useful in gene
CC therapy, and as molecular weight markers for protein electrophoresis
CC purposes. Anti-PRO antibodies may be used in diagnostic assays for
CC PRO polypeptides, or for the affinity purification of the polypeptides
CC from recombinant cell culture or natural sources. ABU69080-ABU69102
CC represent the human PRO polypeptides of the invention.
XX
SQ Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 24; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0

QY 1 MGFLGTGWLVLVLPIDQAPFKPGSQDQKSLHNRELSAERPLNEQTAEAEEDKIKTYPP 60
DB 1 MGFLGTGWLVLVLPIDQAPFKPGSQDQKSLHNRELSAERPLNEQTAEAEEDKIKTYPP 60

QY 61 ENKPGSNYSFVDNLNLKAI TEKEKIEKERSIRSSPLDNKLNVEDVSTKRNKLIDDY 120
DB 61 ENKPGSNYSFVDNLNLKAI TEKEKIEKERSIRSSPLDNKLNVEDVSTKRNKLIDDY 120

QY 121 DSTKSGLDHGFODDPDGLHOLDGTPLTAEDEVHKIAARIYEENDRAVDFKIVSKLINLGL 180
DB 121 DSTKSGLDHGFODDPDGLHOLDGTPLTAEDEVHKIAARIYEENDRAVDFKIVSKLINLGL 180

QY 181 ITESQAHTLEDEVAEVLQKLISK EANNVEEDPNKPTSWTENQAGKIPKVTPMAAIQDGL 240
DB 181 ITESQAHTLEDEVAEVLQKLISK EANNVEEDPNKPTSWTENQAGKIPKVTPMAAIQDGL 240

QY 241 AKGENDETVSNLTLTNGLERTKTYSEDNFEELQVFPNFPYALLKSIDSEKAKEKETLI 300
DB 241 AKGENDETVSNLTLTNGLERTKTYSEDNFEELQVFPNFPYALLKSIDSEKAKEKETLI 300

QY 301 TIMKTLIDFVKMWVKYGTISPEEGSVYLENDDEMIALQTKNLEKATNISKLPAPSE 360
DB 301 TIMKTLIDFVKMWVKYGTISPEEGSVYLENDDEMIALQTKNLEKATNISKLPAPSE 360

QY 361 KSHEETDSTKEAAKMEKEYGSLKSDTKDONSFGKTDPEPKGTEAYLEARKNIWLK 420
DB 361 KSHEETDSTKEAAKMEKEYGSLKSDTKDONSFGKTDPEPKGTEAYLEARKNIWLK 420

QY 421 KHDKKGKEDYDLSMRDFFINKQADAVEKGILDKEEAPAIKRIYSSL 468
DB 421 KHDKKGKEDYDLSMRDFFINKQADAVEKGILDKEEAPAIKRIYSSL 468

RESULT 6
ABU69120
ID ABU69120 standard; Protein; 468 AA.
XX
XX ABU69120;
XX
XX Human PRO polypeptide #18.
XX
XX Human; secreted and transmembrane protein; bone disorder; obesity;
XX cartilage disorder; sports injury; arthritis; diabetes mellitus;
XX hypo-insulinaemia; obesity; hyper-insulinaemia; thalassaemia;
XX haemoglobin-associated disorder; kidney disease; Berger disease;
XX mesangial cell function; nephropathy; Schonlein-Henoch purpura;
XX celiac disease; dermatitis herpetiformis; Crohn's disease;
XX anorectic; antiarthritis; antidiabetic; antianaemic; nephrotropic;
XX antiinflammatory.

```

CC mellitus, hypo-insulinaemia), obesity, hyper-insulinaemia.
CC haemoglobin-associated disorders (e.g. thalassemias), kidney disorders
CC associated with decreased mesangial cell function (e.g. Berger disease),
CC or other nephropathies associated with Schönlein-Henoch purpura,
CC celiac disease, dermatitis herpetiformis or Crohn's disease. The PRO
CC polynucleotide sequences may be used as hybridisation probes in
CC chromosome and gene mapping, or in generating antisense RNA and DNA.
CC They are also useful in preparing PRO polypeptides, in assays to
CC identify other proteins or molecules involved in binding reaction, to
CC generate transgenic animals or knockout animals, which in turn are
CC useful in the development and screening of therapeutically useful
CC reagents, for chromosome identification, and tissue typing. The PRO
CC polypeptides and nucleic acid molecules are also useful in gene
CC therapy, and as molecular weight markers for protein electrophoresis
CC purposes. Anti-PRO antibodies may be used in diagnostic assays for
CC PRO polypeptides, or for the affinity purification of the polypeptides
CC from recombinant cell culture or natural sources. ABU69103-ABU69125
CC represent the human PRO polypeptides of the invention.
XX
SQ Sequence 468 AA;
Query Match 100.0%; Score 2399; DB 24; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGFLGTGFWILVLPVLPQAFPKPGSQDKSLHNLRELGAERPLNQLAEAEEDKIKKTYPP 60
Db 1 MGFLGTGFWILVLPVLPQAFPKPGSQDKSLHNLRELGAERPLNQLAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVDNLNLLKATEKEKTEKQSRSSPLDNKLNVEDVSTKPKRLDDY 120
Db 61 ENKPGQSNYSFVDNLNLLKATEKEKTEKQSRSSPLDNKLNVEDVSTKPKRLDDY 120
Qy 121 DSTKSGLDHKFQDDPDGLHQLDGLPTLTAEDIVHKIAARIYEENDRAVDFDKIVSKLNLGL 180
Db 121 DSTKSGLDHKFQDDPDGLHQLDGLPTLTAEDIVHKIAARIYEENDRAVDFDKIVSKLNLGL 180
Qy 181 ITESQHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAQKIPEKVTPMMAIQDGL 240
Db 181 ITESQHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAQKIPEKVTPMMAIQDGL 240
Qy 241 AKGENDETIVSNTLTITNGLERTTKYSEDNFEELQYFPNFYALLKSIDSEKAEKETLI 300
Db 241 AKGENDETIVSNTLTITNGLERTTKYSEDNFEELQYFPNFYALLKSIDSEKAEKETLI 300
Qy 301 TIMKTLIDFVQWVKYGTISPSEGVSYLENDEMIALQTNKLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVQWVKYGTISPSEGVSYLENDEMIALQTNKLEKNATDNISKLPAPSE 360
Qy 361 KSHEETDSTKEEAARKEYSGLKSDTKDDNSNPGGKTPDEPKGKTEAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEEAARKEYSGLKSDTKDDNSNPGGKTPDEPKGKTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKKNEDYDLKMRDFINKQADAYVEKGIIDKEEAIAIKRIYSSL 468
Db 421 KHKKGKKNEDYDLKMRDFINKQADAYVEKGIIDKEEAIAIKRIYSSL 468

RESULT 7

ABU71580
ID ABU71580 standard; Protein; 468 AA.
XX AC
XX ABU71580;
DT 10-JUN-2003 (first entry)
XX DE Human secreted polypeptide PR05990.
XX KW Human; gene therapy; tumour; cancer.
XX OS Homo sapiens.
XX PN US2003013855-A1.

XX 16-JAN-2003.
XX 03-MAY-2002; 2002US-0063616.
XX 30-DEC-1998; 98KE-0062142.
XX 08-MAR-1999; 99WO-US05028.
XX 14-MAY-1999; 99WO-US10733.
XX 30-DEC-1999; 99WO-US31274.
XX 18-FEB-2000; 2000WO-US04341.
XX 01-MAR-2000; 2000WO-US05601.
XX 02-MAR-2000; 2000WO-US05841.
XX 21-MAR-2000; 2000WO-US07532.
XX 22-MAY-2000; 2000WO-US14042.
XX 02-JUN-2000; 2000WO-US15264.
XX 24-AUG-2000; 2000WO-US23328.
XX 10-NOV-2000; 2000WO-US30873.
XX 01-DEC-2000; 2000WO-US32678.
XX 28-DEC-2000; 2000WO-US34956.
XX 28-FEB-2001; 2001WO-US06520.
XX 01-JUN-2001; 2001WO-US17800.
XX 14-MAY-1999; 99US-0311832.
XX 25-AUG-1999; 99US-0380137.
XX 25-AUG-1999; 99US-0380138.
XX 25-AUG-1999; 99US-0380139.
XX 25-AUG-1999; 99US-0380142.
XX 15-SEP-1999; 99US-0397342.
XX 18-OCT-1999; 99US-0403297.
XX 12-NOV-1999; 99US-0423844.
XX 22-AUG-2000; 2000US-0644848.
XX 18-SEP-2000; 2000US-0664610.
XX 18-SEP-2000; 2000US-0665350.
XX 08-NOV-2000; 2000US-0709238.
XX 20-DEC-2000; 2000US-0747259.
XX 22-MAR-2001; 2001US-0816744.
XX 10-MAY-2001; 2001US-0854208.
XX 10-MAY-2001; 2001US-0854280.
XX 30-MAY-2001; 2001US-0870574.
XX 05-JUN-2001; 2001US-0874503.
XX 29-JUN-2001; 2001US-0869599.
XX 18-JUL-2001; 2001US-0908827.
XX 06-DEC-2001; 2001US-0006867.

(GETH) GENENTECH INC.

Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;

WPI; 2003-330485/31.

N-PSDB; ACAS8884.

New isolated antibody specifically binding a PRO polypeptide, useful
for the preparation of a medicament for treating disorders with the
aberrant expression or activity of the PRO polypeptide, such as tumor
conditions and cancer

Example 17; Page 218-219; 406pp; English.

The invention relates to an antibody that binds to a polypeptide with a
fully defined sequence given in the specification. The methods and
compositions (containing antibodies that specifically bind a PRO
polypeptide) of the present invention are useful for the preparation of a
medicament for the treatment of disorders associated with the aberrant
expression or activity of the PRO polypeptide, such as tumour conditions
and cancer. They can also be used to generate transgenic or knockout
animals useful in the development and screening of therapeutically useful
reagents. The PRO polypeptides and encoding nucleic acids can be used as
molecular weight markers for protein electrophoresis, chromosome
identification and tissue typing. The PRO polypeptides are useful to
induce angiogenesis e.g wound healing; in the treatment of sports-related
joint problems, articular cartilage defects, osteoarthritis or rheumatoid
arthritis; diabetes; hyperinsulinaemia and hypoinsulinaemia. The
antibodies may be used in various diagnostic, competitive binding and/or

PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 22-FEB-2000; 2000WO-US04342.
 PR 24-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 01-MAR-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05746.
 PR 10-MAR-2000; 2000WO-US05841.
 PR 10-MAR-2000; 2000WO-US06319.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 11-AUG-2000; 2000WO-US22031.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-MAR-2001; 2001WO-US06666.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 22-JUN-2001; 2001WO-US20116.
 PR 23-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 20-DEC-2000; 2000US-0747259.
 PR 28-FEB-2001; 2001US-0796498.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0808689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 18-MAY-2001; 2001US-0854280.
 PR 25-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 01-JUN-2001; 2001US-0866034.
 PR 05-JUN-2001; 2001US-0872035.
 PR 14-JUN-2001; 2001US-0874503.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.
 (GETH) GENENTECH INC.

XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 DR WPI: 2003-332040/31.
 DR N-PSDB; ACA03647.

XX New secreted and transmembrane PRO nucleic acids, useful for gene
 PT therapy, in chromosome and gene mapping, as chromosome markers, in
 PT tissue typing, and in chromosome identification

XX Claim 12; Fig 90; 660pp; English.

XX The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The

CC PRO polypeptides are secreted and transmembrane proteins. The PRO
 CC polypeptides are useful for detecting other PRO polypeptides, for
 CC linking bioactive molecules to cells expressing PRO polypeptides,
 CC for modulating biological activities of cells expressing PRO
 CC polypeptides, and for identifying agonists or antagonists.
 CC The PRO polypeptides are useful for stimulating the release of
 CC tumour necrosis factor (TNF)-alpha from human blood, for stimulating the
 CC proliferation or differentiation of chondrocytes, and detecting the
 CC presence of tumours. The polynucleotide sequences encoding PRO
 CC polypeptides are useful as hybridisation probes, in chromosome and
 CC gene mapping, in the generation of antisense RNA and DNA, in the
 CC preparation of PRO polypeptides, for generating transgenic animals or
 CC knockout animals, for the genetic analysis of individuals with genetic
 CC disorders, and in gene therapy. ABU66570-ABU66844 represent the human
 CC PRO polypeptides of the invention.
 CC Note: the sequence data for this patent was obtained in electronic
 CC format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipdsIDEntry.html.
 XX

SQ Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 24; Length 468;
 Best Local Similarity 100.0%; Pred. No. 4.1e-146;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGFLGTGWLVLVLPQAPKPGSQDKSLHNRLSAERPLNEQIAEAEEDKIKKTYPP 60
 DB 1 MGFLGTGWLVLVLPQAPKPGSQDKSLHNRLSAERPLNEQIAEAEEDKIKKTYPP 60
 QY 61 ENKPGQSNYSFVDNLNLLKAI TEKEKIEKERSIRSSPLDNKLVNVEDVSTNKRKLIDDY 120
 DB 61 ENKPGQSNYSFVDNLNLLKAI TEKEKIEKERSIRSSPLDNKLVNVEDVSTNKRKLIDDY 120
 QY 121 DSTKSGLDHKEODDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVPDKIVSKLLNLGL 180
 DB 121 DSTKSGLDHKEODDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVPDKIVSKLLNLGL 180
 QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENOAGKIPEKVTMPMAIQDGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENOAGKIPEKVTMPMAIQDGL 240
 QY 241 AKGENDETVSNTLTITNGLERRTKTSYSEDNPEELQYFPNFVALLKSIDSEKAKEKETLI 300
 DB 241 AKGENDETVSNTLTITNGLERRTKTSYSEDNPEELQYFPNFVALLKSIDSEKAKEKETLI 300
 QY 301 TIMKTLIDFVQWVKYGTISPSEGVSYLENDEMIALQTNKLEKNATDNISKLPAPSE 360
 DB 301 TIMKTLIDFVQWVKYGTISPSEGVSYLENDEMIALQTNKLEKNATDNISKLPAPSE 360
 QY 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDONSPPGKTDPEKGTAYLEAIRKNIEWLK 420
 DB 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDONSPPGKTDPEKGTAYLEAIRKNIEWLK 420
 QY 421 KHDKKGKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
 DB 421 KHDKKGKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468

RESULT 11

ABU66890
 ID ABU66890 standard; Protein; 468 AA.

XX AC
 XX ABU66890;

XX 27-MAY-2003 (first entry)

XX Human secreted/transmembrane, PRO, protein SEQ ID 90.

XX Human; secreted protein; transmembrane protein; PRO;
 KW inflammatory disease; organ failure; atherosclerosis; cardiac injury;
 KW infertility; birth defects; premature aging; AIDS; biosensor;
 KW acquired immunodeficiency syndrome; cancer; diabetic complication;
 KW bioreactor; tumour.

XX OS Homo sapiens.
 XX PN US2003032155-A1.
 XX PD 13-FEB-2003.
 XX PF 03-MAY-2002; 2002US-0137865.
 XX PR 31-MAR-1997; 97WO-US05230.
 PR 12-JUN-1998; 98WO-US12456.
 PR 14-JUL-1998; 98WO-US14552.
 PR 28-AUG-1998; 98WO-US17888.
 PR 10-SEP-1998; 98WO-US18824.
 PR 14-SEP-1998; 98WO-US19093.
 PR 14-SEP-1998; 98WO-US19094.
 PR 14-SEP-1998; 98WO-US19177.
 PR 16-SEP-1998; 98WO-US19330.
 PR 17-SEP-1998; 98WO-US19437.
 PR 07-OCT-1998; 98WO-US21141.
 PR 29-OCT-1998; 98WO-US22991.
 PR 29-OCT-1998; 98WO-US22992.
 PR 20-NOV-1998; 98WO-US24855.
 PR 01-DEC-1998; 98WO-US25108.
 PR 05-JAN-1999; 99WO-US00106.
 PR 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99WO-US05190.
 PR 20-APR-1999; 99WO-US08615.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 20-DEC-1999; 99WO-US30999.
 PR 22-DEC-1999; 99WO-US30720.
 PR 30-DEC-1999; 99WO-US31243.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 06-JAN-2000; 2000WO-US00376.
 PR 11-FEB-2000; 2000WO-US03565.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05746.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 10-MAR-2000; 2000WO-US06319.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 11-AUG-2000; 2000WO-US22031.

PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-MAR-2001; 2001WO-US06666.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 22-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US20116.
 PR 09-JUL-2001; 2001WO-US21086.
 PR 20-DEC-2000; 2000US-0747259.
 PR 28-FEB-2001; 2001US-0796498.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0808689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 18-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 25-MAY-2001; 2001US-0866034.
 PR 01-JUN-2001; 2001US-0872035.
 PR 05-JUN-2001; 2001US-0874503.
 PR 14-JUN-2001; 2001US-0882636.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.

(GETH) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 WPI; 2003-331925/31.
 N-PSDB; ACA04068.

New secreted and transmembrane nucleic acids and polypeptides.
 designated as PRO, useful for treating inflammation, organ failure,
 atherosclerosis, cardiac injury, infertility, birth defects, premature
 aging, AIDS, or cancer

Claim 12; Fig 90; 659pp; English.

The invention relates to an isolated nucleic acid comprising, or which is
 at least 80% identical to, or the full-length coding sequence of, any of
 the 275 nucleotide sequences, encoding the corresponding PRO polypeptide
 (one of 275 secreted or transmembrane proteins). The nucleic acid
 further comprises the full-length coding sequence of the DNA deposited
 under American Type Culture Collection (ATCC) accession number in a list
 given in the specification. Also included are vectors and host
 cells for producing PRO proteins, PRO fusion proteins, anti-PRO
 antibodies, PRO extracellular domains and mature sequences, methods
 of detecting PRO proteins, methods for stimulating the release of
 TNF-alpha (tumour necrosis factor alpha) from human blood,
 and the proliferation of differentiation of chondrocyte cells, the
 proliferation of, or gene expression in pericyte cells, the release or
 proteoglycans from cartilage, proliferation of inner ear utricular
 supporting cells, the proliferation of T-lymphocyte cells, the release
 of a cytokine from peripheral blood mononuclear cells (PBMC), or the
 proliferation of endothelial cells), a method for modulating the uptake
 of glucose or free fatty acid (FFA) by skeletal muscle cells,
 a method for inhibiting the binding of A-peptide to factor VIIA,
 or the differentiation of adipocyte cells, a method for detecting the
 presence of a tumour in a mammal and an oligonucleotide probe derived

CC from any of the nucleotide sequences cited above. The nucleic acids and
 CC polypeptides are useful for treating inflammatory diseases, organ
 CC failure, atherosclerosis, cardiac injury, infertility, birth defects,
 CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or
 CC diabetic complications. The nucleic acids are useful as hybridisation
 CC probes, in chromosome and gene mapping, and in generating antisense RNA
 CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,
 CC biosensors or bioreactors. Both are useful in tissue typing.
 CC The present sequence represents a PRO protein of the invention.
 XX
 SQ Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 24; Length 468;
 Best Local Similarity 100.0%; Pred. No. 4.1e-146;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGFLGTGWILVLPQAPKPGSGDSKSLHRELGAERPLNEQIAEAEDKIKKTYPP 60
 DB 1 MGFLGTGWILVLPQAPKPGSGDSKSLHRELGAERPLNEQIAEAEDKIKKTYPP 60
 QY 61 ENKFGQSNYSFVNLLNLKATTEKEKIEKQSRSSPLDNKLNVEDVDSTKRNKLIDDY 120
 DB 61 ENKFGQSNYSFVNLLNLKATTEKEKIEKQSRSSPLDNKLNVEDVDSTKRNKLIDDY 120
 QY 121 DSTKSGLDHFKQDDPDGHLQDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGL 180
 DB 121 DSTKSGLDHFKQDDPDGHLQDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGL 180
 QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKPEKVTPTMAAQDGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKPEKVTPTMAAQDGL 240
 QY 241 AKGENDTVSNLTLTNGLERRTKTSYSEDNFEELQYFPNFYALLKSIDSEKAEKETLI 300
 DB 241 AKGENDTVSNLTLTNGLERRTKTSYSEDNFEELQYFPNFYALLKSIDSEKAEKETLI 300
 QY 301 TIMKTLIDFVKMVKYGTISPEEGSVYLENDLMIALQTNKLEKATDNISKLPFAPSE 360
 DB 301 TIMKTLIDFVKMVKYGTISPEEGSVYLENDLMIALQTNKLEKATDNISKLPFAPSE 360
 QY 361 KSHEETDSTKEEAARKEKEYGSLKSDNSPGGKTDPEKGTAYLEAIRKNIEWLK 420
 DB 361 KSHEETDSTKEEAARKEKEYGSLKSDNSPGGKTDPEKGTAYLEAIRKNIEWLK 420
 QY 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGILDKGEAEAIKRIYSSL 468
 DB 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGILDKGEAEAIKRIYSSL 468

RESULT 12
 ABUS9695
 ID ABUS9695 standard; Protein; 468 AA.
 XX
 AC ABUS9695;
 XX
 DT 13-MAY-2003 (first entry)
 XX
 DE Novel secreted and transmembrane protein PRO5990.
 XX
 KW Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
 KW cardiac insufficiency disorder; cancer; tumour; immune response;
 KW adrenal cortical capillary endothelial growth; c-fos induction;
 KW vascular endothelial growth factor inhibition; VEGF inhibition;
 KW endothelial cell growth inhibitor; T-lymphocytes stimulation;
 KW retinal neurons cell survival; rod photoreceptor cell survival;
 KW retinal disorder; retinitis pigmentosa; kidney disorder;
 KW mammalian kidney mesangial cell proliferation; Berger disease;
 KW dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
 KW chondrocyte redifferentiation; sports injury; arthritis.
 XX
 OS Homo sapiens.
 XX
 PN US2003017563-A1.

XX
 PD 23-JAN-2003.
 XX
 XX 07-MAY-2002; 2002US-0140808.
 XX
 XX 31-MAR-1997; 97WO-US05230.
 XX 12-JUN-1998; 98WO-US12456.
 XX 14-JUL-1998; 98WO-US14552.
 XX 28-AUG-1998; 98WO-US17888.
 XX 10-SEP-1998; 98WO-US18824.
 XX 14-SEP-1998; 98WO-US19093.
 XX 14-SEP-1998; 98WO-US19094.
 XX 14-SEP-1998; 98WO-US19177.
 XX 16-SEP-1998; 98WO-US19330.
 XX 17-SEP-1998; 98WO-US19437.
 XX 07-OCT-1998; 98WO-US21141.
 XX 29-OCT-1998; 98WO-US22991.
 XX 29-OCT-1998; 98WO-US22992.
 XX 20-NOV-1998; 98WO-US24855.
 XX 01-DEC-1998; 98WO-US25108.
 XX 05-JAN-1999; 99WO-US00106.
 XX 08-MAR-1999; 99WO-US05028.
 XX 10-MAR-1999; 99WO-US05190.
 XX 20-APR-1999; 99WO-US08615.
 XX 14-MAY-1999; 99WO-US10733.
 XX 01-SEP-1999; 99WO-US20111.
 XX 08-SEP-1999; 99WO-US20594.
 XX 13-SEP-1999; 99WO-US20944.
 XX 15-SEP-1999; 99WO-US21090.
 XX 15-SEP-1999; 99WO-US21547.
 XX 05-OCT-1999; 99WO-US22089.
 XX 29-NOV-1999; 99WO-US28214.
 XX 30-NOV-1999; 99WO-US28313.
 XX 30-NOV-1999; 99WO-US28409.
 XX 01-DEC-1999; 99WO-US28301.
 XX 01-DEC-1999; 99WO-US28634.
 XX 02-DEC-1999; 99WO-US28551.
 XX 02-DEC-1999; 99WO-US28564.
 XX 02-DEC-1999; 99WO-US28565.
 XX 16-DEC-1999; 99WO-US30095.
 XX 20-DEC-1999; 99WO-US30911.
 XX 20-DEC-1999; 99WO-US30999.
 XX 22-DEC-1999; 99WO-US30720.
 XX 30-DEC-1999; 99WO-US31243.
 XX 30-DEC-1999; 99WO-US31274.
 XX 05-JAN-2000; 2000WO-US00219.
 XX 06-JAN-2000; 2000WO-US00277.
 XX 06-JAN-2000; 2000WO-US00376.
 XX 11-FEB-2000; 2000WO-US03565.
 XX 18-FEB-2000; 2000WO-US04341.
 XX 18-FEB-2000; 2000WO-US04342.
 XX 22-FEB-2000; 2000WO-US04414.
 XX 24-FEB-2000; 2000WO-US04914.
 XX 24-FEB-2000; 2000WO-US05004.
 XX 01-MAR-2000; 2000WO-US05601.
 XX 02-MAR-2000; 2000WO-US05746.
 XX 02-MAR-2000; 2000WO-US05841.
 XX 10-MAR-2000; 2000WO-US06319.
 XX 15-MAR-2000; 2000WO-US06884.
 XX 20-MAR-2000; 2000WO-US07377.
 XX 21-MAR-2000; 2000WO-US07532.
 XX 30-MAR-2000; 2000WO-US08439.
 XX 17-MAY-2000; 2000WO-US13705.
 XX 22-MAY-2000; 2000WO-US14042.
 XX 30-MAY-2000; 2000WO-US14941.
 XX 02-JUN-2000; 2000WO-US15264.
 XX 28-JUL-2000; 2000WO-US20710.
 XX 11-AUG-2000; 2000WO-US22031.
 XX 23-AUG-2000; 2000WO-US23522.
 XX 24-AUG-2000; 2000WO-US23328.
 XX 08-NOV-2000; 2000WO-US30952.
 XX 10-NOV-2000; 2000WO-US30873.

01-DEC-2000; 2000WO-US32678.
 28-DEC-2000; 2000WO-US34956.
 28-FEB-2001; 2001WO-US06520.
 01-MAR-2001; 2001WO-US06656.
 25-MAY-2001; 2001WO-US17092.
 01-JUN-2001; 2001WO-US17800.
 20-JUN-2001; 2001WO-US19692.
 22-JUN-2001; 2001WO-US20116.
 29-JUN-2001; 2001WO-US21066.
 09-JUL-2001; 2001WO-US21735.
 20-DEC-2000; 2000US-074259.
 28-FEB-2001; 2001US-0796498.
 09-MAR-2001; 2001US-0802706.
 14-MAR-2001; 2001US-0808689.
 22-MAR-2001; 2001US-0816744.
 05-APR-2001; 2001US-0828366.
 10-MAY-2001; 2001US-0854208.
 10-MAY-2001; 2001US-0854280.
 18-MAY-2001; 2001US-0860216.
 25-MAY-2001; 2001US-0866028.
 25-MAY-2001; 2001US-0866034.
 01-JUN-2001; 2001US-0872035.
 05-JUN-2001; 2001US-0874503.
 14-JUN-2001; 2001US-0882636.
 19-JUN-2001; 2001US-0886342.
 21-JUN-2001; 2001US-0887879.
 18-JUL-2001; 2001US-0908827.
 06-AUG-2001; 2001US-0924419.
 09-AUG-2001; 2001US-0927796.
 16-AUG-2001; 2001US-0931836.
 19-DEC-2001; 2001US-0028072.
 (GETH) GENENTECH INC.
 Baker KP, Beresini M, Deforge L, Desnoyers L, Filvaroff E, Gao W;
 Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 WPI; 2003-148238/14.
 N-PSDB; AEX89185.
 Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
 and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
 are therapeutically useful for enhancing immune response and in cancer
 treatments -
 Claim 12; Fig 90; 659pp; English.
 The invention describes an isolated human PRO polypeptide. The PRO
 polypeptides are useful in detecting PRO polypeptides in a sample, in
 linking a bioactive molecule to a cell expressing a PRO polypeptide, and
 in modulating at least one biological activity of a cell expressing a PRO
 polypeptide. PRO132 stimulates hypertrophy of neonatal heart and is thus
 useful for treating cardiac insufficiency disorders. PRO1154 and PRO1186
 stimulate adrenal cortical capillary endothelial growth, and PRO536,
 PRO943, PRO828, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,
 PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus
 useful for treating conditions or disorders where angiogenesis would be
 beneficial, e.g. wound healing and antagonist of this polypeptide are
 useful for treating cancerous tumours. PRO812 inhibits vascular
 endothelial growth factor (VEGF) stimulated proliferation of endothelial
 cells and is thus useful for inhibiting endothelial cell growth in
 mammals which would be beneficial in inhibiting tumour growth. PRO826,
 PRO1068, PRO1184, and PRO1346 stimulate proliferation of
 stimulated T-lymphocytes and are therapeutically useful for enhancing
 immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of
 retinal neurons cells (PRO1132 is also enhances survival/proliferation of
 rod photoreceptor cells) and therefore are useful for treating retinal
 disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813
 and PRO1066 induce proliferation of mammalian kidney mesangial cells,
 and therefore are useful for treating kidney disorders associated with
 decreased mesangial cell function such as Berger disease or other
 nephropathies associated with dermatitis, herpeticiformis or Crohn's

CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
 CC proliferation and/or redifferentiation of chondrocytes in culture and
 CC are thus useful for treating sports injuries, and arthritis. This
 CC is the amino acid sequence of a novel human PRO protein.
 XX
 SQ Sequence 468 AA;
 Query Match 100.0%; Score 2399; DB 24; Length 468;
 Best Local Similarity 100.0%; Pred. No. 4,1e-146;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGFLGTGTLVLVLPLOAPFKPGSGQKSLHNRLELSAEPLNEQIAEAEEDKIKKTYPP 60
 DB 1 MGFLGTGTLVLVLPLOAPFKPGSGQKSLHNRLELSAEPLNEQIAEAEEDKIKKTYPP 60
 QY 61 ENKPGQSNYSFVDNLLNLLKAI TEKEIEKERQIRSPDLNKLNVEDVSTKRNKLIDDDY 120
 DB 61 ENKPGQSNYSFVDNLLNLLKAI TEKEIEKERQIRSPDLNKLNVEDVSTKRNKLIDDDY 120
 QY 121 DSTKSGLDHFKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLMLGL 180
 DB 121 DSTKSGLDHFKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLMLGL 180
 QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTWMAAIDQGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTWMAAIDQGL 240
 QY 241 AKGENDETWSNTLTNLGLERRTKTYSEDNFELQYFPNFYALLKSIDSEKAKEKETLI 300
 DB 241 AKGENDETWSNTLTNLGLERRTKTYSEDNFELQYFPNFYALLKSIDSEKAKEKETLI 300
 QY 301 TIMKTLLIDFVKMVKYGTISPEEGSVYLENDEMIALQTKNKLKKNATDNISKLPAPSE 360
 DB 301 TIMKTLLIDFVKMVKYGTISPEEGSVYLENDEMIALQTKNKLKKNATDNISKLPAPSE 360
 QY 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDSDNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
 DB 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDSDNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
 QY 421 KHDKGNKEDYDLSKMRDFINKQADAVVEKGILDKEEAIAIKRIYSSL 468
 DB 421 KHDKGNKEDYDLSKMRDFINKQADAVVEKGILDKEEAIAIKRIYSSL 468
 RESULT 13
 ABG73287
 ID ABG73287 standard; Protein; 468 AA.
 XX AC ABG73287;
 XX AC
 XX DT 28-APR-2003 (first entry)
 XX DE Human neurosecretory protein-1 (HUNSP-1).
 KW Human; neurosecretory protein-1; HUNSP-1; neurological disorder;
 KW epilepsy; Huntington's disease; dementia; Alzheimer's disease;
 KW multiple sclerosis; Parkinson's disease; schizophrenia; amnesia;
 KW endocrine disorder; hypogonadism; Sheehan's syndrome; leukaemia;
 KW Hashimoto's disease; Grave's disease; Cushing's disease; cancer;
 KW adenocarcinoma; diabetes insipidus; melanoma; myeloma sarcoma;
 KW bone cancer; breast cancer; cancer of the pancreas; brain cancer;
 KW prostate cancer; skin cancer; thyroid cancer; antidiabetic; neurotropic;
 KW neuroprotective; neuroleptic; cytostatic; anticonvulsant; thyromimetic;
 KW lymphoma; antiparkinsonian; antithyroid; immunosuppressive; antithyroid;
 KW pancreatic islet cell cdna library; ISUTNOT01; Incyte clone 2379427.
 OS Homo sapiens.
 XX OS
 PN US2002156014-A1.
 XX PD 24-OCT-2002.
 XX PF 14-FEB-2002; 2002US-0077040.

XX PR 17-APR-1998; 98US-0062601.
XX PA (INCY-) INCYTE PHARM INC.
XX PI Lal P, Corley NC, Patterson C;
XX DR WPI; 2003-238117/23.
XX DR N-PSDB; ABX11115.
XX PT New human neurosecretory proteins (HUNSP) and polynucleotides, useful
XX PT for diagnosing, treating or preventing e.g. epilepsy, Alzheimer's
XX PT disease, schizophrenia, Grave's disease, leukaemia, lymphoma, or breast
XX PT or brain cancer
XX PS Claim 1; Fig 1; 45pp; English.
XX CC The present invention relates to the isolation of human neurosecretory
XX CC proteins -1 and -2 (HUNSP-1 and HUNSP-2), and the polynucleotide
XX CC sequences encoding them. The HUNSP polypeptides, polynucleotides, and
XX CC agonists are useful for treating a disease or condition associated with
XX CC decreased expression of functional HUNSP. Antagonists of HUNSP are
XX CC useful for treating a disease or condition associated with
XX CC overexpression of functional HUNSP. HUNSP antibodies are useful for
XX CC diagnosing a condition or disease associated with the expression of
XX CC HUNSP. The sequences of the invention are particularly useful for
XX CC diagnosing, treating or preventing neurological disorders
XX CC (e.g. epilepsy, Huntington's disease, dementia, Alzheimer's disease,
XX CC multiple sclerosis, Parkinson's disease, schizophrenia, amnesia),
XX CC endocrine disorders (e.g. hypogonadism, Sheehan's syndrome, diabetes
XX CC insipidus, Hashimoto's disease, Grave's disease, Cushing's disease),
XX CC and cancers (e.g. adenocarcinoma, leukaemia, lymphoma, melanoma,
XX CC myeloma sarcoma, bone cancer, breast cancer, cancer of the pancreas,
XX CC brain cancer, prostate cancer, skin cancer or thyroid cancer).
XX CC The present sequence represents HUNSP-1. The cDNA encoding HUNSP-1
XX CC is isolated from pancreatic islet cell cDNA library (ISLNOT01) Incyte
XX CC clone 2379427.
XX SQ Sequence 468 AA;
Query Match 100.0%; Score 2399; DB 24; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146; Indels 0; Gaps 0;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGTWILVLPIQAFPKPGSQDLSHNRSLSAERPLNEQIAEAEEDKIKKTYPP 60
DB 1 MGFLGTGTWILVLPIQAFPKPGSQDLSHNRSLSAERPLNEQIAEAEEDKIKKTYPP 60
QY 61 ENKPGQSNYSFVDNLLKAITKEKIEKQIRSRSPDLNKLNVEDVSTKNRKLIDY 120
DB 61 ENKPGQSNYSFVDNLLKAITKEKIEKQIRSRSPDLNKLNVEDVSTKNRKLIDY 120
QY 121 DSTKSGLDHFKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180
DB 121 DSTKSGLDHFKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180
QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTWMAAQDGL 240
DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTWMAAQDGL 240
QY 241 AKGENDETUNTLTNGLRRTKTYSEDNFELOQFPNFPYALLKSIDSEKAKEKETLI 300
DB 241 AKGENDETUNTLTNGLRRTKTYSEDNFELOQFPNFPYALLKSIDSEKAKEKETLI 300
QY 301 TINKTLIDFVGMVVKGTISPEGVSVLENLDMIALQTKNLEKNATNISKLPAPSE 360
DB 301 TINKTLIDFVGMVVKGTISPEGVSVLENLDMIALQTKNLEKNATNISKLPAPSE 360
QY 361 KSHETDSTKEAAKMEKYGSLKSTKODNSNPGKTDPEKGTKEAYLEAIRKNIWLK 420
DB 361 KSHETDSTKEAAKMEKYGSLKSTKODNSNPGKTDPEKGTKEAYLEAIRKNIWLK 420
QY 421 KHDKKGKEDYDLSKWRDFINKQADAYVEKGILDKEEAIAIKRIYSSL 468

DB 421 KHDKKGKEDYDLSKWRDFINKQADAYVEKGILDKEEAIAIKRIYSSL 468
RESULT 14
AAO13905
ID AAO13905 standard; Protein; 567 AA.
XX AC AAO13905;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide SEQ ID NO 27797.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US04927.
XX PR 28-FEB-2000; 2000US-0515126.
XX PR 18-MAY-2000; 2000US-0577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX DR N-PSDB; AAI93836.
XX PT Isolated nucleic acids and polypeptides, useful for preventing
XX PT diagnosing and treating e.g. leukaemia, inflammation and immune
XX PT disorders -
XX PS Claim 20; SEQ ID NO 27797; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI93841) and
XX CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation.
XX CC Note: The sequence data for this patent did not form part of the printed
XX CC specification, but was obtained in electronic format directly from WIPO
XX CC at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 567 AA;
Query Match 100.0%; Score 2399; DB 22; Length 567;
Best Local Similarity 100.0%; Pred. No. 5.3e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGTWILVLPIQAFPKPGSQDLSHNRSLSAERPLNEQIAEAEEDKIKKTYPP 60
DB 100 MGFLGTGTWILVLPIQAFPKPGSQDLSHNRSLSAERPLNEQIAEAEEDKIKKTYPP 159
QY 61 ENKPGQSNYSFVDNLLKAITKEKIEKQIRSRSPDLNKLNVEDVSTKNRKLIDY 120
DB 160 ENKPGQSNYSFVDNLLKAITKEKIEKQIRSRSPDLNKLNVEDVSTKNRKLIDY 219
QY 121 DSTKSGLDHFKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180

Db 220 DSTKSGLDHFKODDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 279
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
Db 280 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 339
Qy 241 AKGENDETIVSNTLTITNGLERRTKTYSEDNFEELYFPNFVALLKSIDSEKAKEKETLI 300
Db 340 AKGENDETIVSNTLTITNGLERRTKTYSEDNFEELYFPNFVALLKSIDSEKAKEKETLI 399
Qy 301 TIMKTLIDFVKMMVKYGTISPPEGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
Db 400 TIMKTLIDFVKMMVKYGTISPPEGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 459
Qy 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 420
Db 460 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 519
Qy 421 KHDKKGKEDYDLKMDRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
Db 520 KHDKKGKEDYDLKMDRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 567

RESULT 15

AAW93910
ID AAW93910 standard; Protein; 468 AA.

AC AAW93910;

DT 06-NOV-2001 (first entry)

DE Human polypeptide, SEQ ID NO: 4061.

DE Human; full length cDNA; cDNA synthesis; oligo-capping.

OS Homo sapiens.

PN EP1130094-A2.

PD 05-SEP-2001.

PF 07-JUL-2000; 2000EP-0114089.

PR 08-JUL-1999; 99JP-0194486.

PR 11-JAN-2000; 2000JP-0118774.

PR 02-MAY-2000; 2000JP-0183765.

PA (HELI-) HELIX RES INST.

PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;

PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otsuki T, Koga H;

DR WPI; 2001-524255/58.

DR N-PSDB; AAK94872.

PT 830 Primers useful for synthesizing full length cDNA clones and their

PT use in genetic manipulation -

PS Claim 8; SEQ ID NO 4061; 1380pp + sequence listing; English.

CC The invention relates to primers for synthesising full length cDNA
CC clones. 830 cDNA molecules encoding a human protein have been
CC isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA
CC molecules have been determined. Primers for synthesising the full length
CC cDNA are useful for clarifying the function of the protein encoded by
CC the cDNA. The full length clones were obtained by construction of full
CC length enriched cDNA libraries that were synthesised by the oligo-capping
CC method. The primers enable the production of the full length cDNA easily
CC without any special methods. The present sequence is a polypeptide
CC encoded by a full length human cDNA of the invention.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in CD-ROM format directly from EPO.

XX SQ Sequence 468 AA;
Query Match 99.7%; Score 2392; DB 22; Length 468;
Best Local Similarity 99.8%; Pred. No. 1.2e-145;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGFLGTGTTWLVLPLOAPFKPGSQDKSLHNRHLSAERPLNEQIAEABEDKIKKTYPP 60
Db 1 MGFLGTGTTWLVLPLOAPFKPGSQDKSLHNRHLSAERPLNEQIAEABEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVDMNLNLLKAITKEKEIKERQSISSPLDNKLNVEDVDSTKNRKLDDY 120
Db 61 ENKPGQSNYSFVDMNLNLLKAITKEKEIKERQSISSPLDNKLNVEDVDSTKNRKLDDY 120
Qy 121 DSTKSGLDHFKODDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKODDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
Qy 241 AKGENDETIVSNTLTITNGLERRTKTYSEDNFEELYFPNFVALLKSIDSEKAKEKETLI 300
Db 241 AKGENDETIVSNTLTITNGLERRTKTYSEDNFEELYFPNFVALLKSIDSEKAKEKETLI 300
Qy 301 TIMKTLIDFVKMMVKYGTISPPEGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVKMMVKYGTISPPEGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
Qy 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 420
Qy 421 KHDKKGKEDYDLKMDRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
Db 421 KHDKKGKEDYDLKMDRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468

Search completed: January 2, 2004, 12:34:22

Job time : 47 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 12:33:21 ; Search time 21 Seconds
(without alignments)
942.928 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGWLVLVLPQAF.....EKGILDKBAEAKRIYSSL 468

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	179	7.5	1786	3	US-08-973-462-8
2	172.5	7.2	1164	3	US-08-923-992A-2
3	169	7.0	1098	3	US-08-923-992A-8
4	168	7.0	1164	3	US-08-923-992A-10
5	167.5	7.0	957	4	US-09-914-259-16
6	166	6.9	1104	3	US-08-923-992A-4
7	166	6.9	1128	3	US-08-923-992A-6
8	166	6.9	2662	4	US-09-595-684B-31
9	163.5	6.8	956	4	US-09-914-259-17
10	162	6.8	8991	4	US-08-714-741-32
11	159.5	6.6	816	2	US-08-533-306A-6
12	159.5	6.6	816	2	US-08-742-923A-6
13	158.5	6.6	885	2	US-08-533-306A-4
14	158.5	6.6	885	2	US-08-742-923A-4
15	156.5	6.5	1087	4	US-09-914-259-12
16	156	6.5	1354	3	US-08-685-871-2
17	154.5	6.4	1231	4	US-08-714-741-41
18	153.5	6.4	984	1	US-08-714-481-2
19	153.5	6.4	984	5	PCT-US95-06111-2
20	153.5	6.4	984	5	PCT-US93-07261-11
21	153	6.4	1588	5	PCT-US93-07261-16
22	153	6.4	1663	5	PCT-US93-07261-16
23	153	6.4	2285	4	US-09-308-375-2
24	150.5	6.3	1010	4	US-09-134-001C-5178
25	147	6.1	534	4	US-09-103-664A-2
26	146.5	6.1	1388	2	US-08-685-576-1
27	146.5	6.1	1388	2	US-08-685-576-4

28	145.5	6.1	652	4	US-08-559-896B-2	Sequence 2, Appli
29	145	6.0	588	4	US-08-714-741-42	Sequence 42, Appli
30	144.5	6.0	343	3	US-08-937-271-17	Sequence 17, Appli
31	144	6.0	10182	4	US-09-134-001C-3159	Sequence 3159, Ap
32	143.5	6.0	864	4	US-08-714-741-40	Sequence 40, Appli
33	143	6.0	3878	4	US-09-914-259-11	Sequence 11, Appli
34	142	5.9	967	4	US-09-914-259-21	Sequence 21, Appli
35	142	5.9	1032	4	US-09-914-259-26	Sequence 26, Appli
36	141.5	5.9	1146	4	US-08-914-999-6	Sequence 6, Appli
37	141	5.9	1507	3	US-08-929-329-5	Sequence 5, Appli
38	139.5	5.8	963	4	US-09-914-259-20	Sequence 20, Appli
39	139	5.8	1939	4	US-09-310-187A-1	Sequence 1, Appli
40	139	5.8	3111	2	US-08-460-309-4	Sequence 4, Appli
41	139	5.8	3111	2	US-08-125-077-4	Sequence 4, Appli
42	138	5.8	630	3	US-08-973-462-9	Sequence 9, Appli
43	138	5.8	1177	4	US-09-134-001C-5106	Sequence 5106, Ap
44	138	5.8	1234	4	US-09-592-054-8	Sequence 8, Appli
45	137.5	5.7	586	2	US-08-630-822A-70	Sequence 70, Appli

ALIGNMENTS

RESULT 1

US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; TITLE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; FILE REFERENCE: 0660-0125-0 PCT
; CURRENT APPLICATION NUMBER: US/08/973.462B
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide

US-08-973-462-8

Query Match 7.5%; Score 179; DB 3; Length 1786;
Best Local Similarity 21.5%; Pred. No. 0.00012;
Matches 118; Conservative 96; Mismatches 197; Indels 138; Gaps 25;

Qy	27	QDKSLHNRISAE--RPLNEQIAEAEEDKIKTYPPENKPGQSN-----YSFVDNL-- 75
Db	957	EENAVESNENVAENLEKLNVTNVLDKVEETVEISGESLENNEMDKAFSEIFDNVKG 1016
Qy	76	---NLLK-----AITEKEKIEROSIRSSPLDNKLN-----VEDVDSTK-- 112
Db	1017	IOENLLTGMFRSTIETSIQSEKVDI--NENVSSILDNIENKKEGLLNKLENISSTEGV 1075
Qy	113	-----NRLKIDDDYD-----STKGGLDHKFDQDDPGLHQLDGTPLTAEI-- 151
Db	1076	QETVTEHVEQNVYVDVDPAMKQDFLGILNEAGLKEMFNFLED--VPKSESDVITVEIK 1134
Qy	152	-----VHKIARIYEENDRAVDFKIVSKLLNGLITESQANTLEDEVAEVLQKLSKE 204
Db	1135	DEPVQKEVEKETVSIIEEMEENIVD-----VLEEKEDLTOKMDAVESEIIS 1183
Qy	205	ANNYEDPNKPTSTWENOAGKIPKVTPTMAAIQDLAKGENDETSTNTLTITNGLERRTK 264
Db	1184	SDSKEE-----TES-----IKDKKDVSLVVEEVDNDDESVEKVELKNEEBELMK 1231
Qy	265	TYSEDN-----FEELQYFFNFYA--LLKSIDSEKAK-----EKETLITIMK 304

Db 1232 DAVEINDITSLIBETQELNEVEADLIKOMEKLEKALSEDSKEIIDAKDDTLEKIE 1291
 QY 305 -----TLIDFVKMVKYGTISPE--EGSVYLENDIMIALQTK--NLEKNATDNISKL 354
 Db 1292 EHDITTTILDEV---VELKQVEDKIEKVSOLKLEEDILKEVKEIKELSEILEDYKEL 1348
 QY 355 FPAPSEKSHETDSTKEEAAKMEKEYGSLKDKSTDDNSNPGKTDPEPKGKTEAYLEAIRK 414
 Db 1349 KTIETDILEEKEIEKDHFKFEFEAEIEKOLEADILKEVSVLESEBEEKLEEVHE-LKE 1407
 QY 415 NLEWLKHDK--KGNKEDYDLSKMRDFINKQADAY---VEKGILDK-----EEA 458
 Db 1408 EVEHIIISDAHIGKLEED-DLEEVDLKGSLDMLKGMELGMDMKESLEDVTTTKLGERV 1466
 QY 459 EAIKRIYSS 467
 Db 1467 ESKDVLSS 1475

RESULT 2

US-08-923-992A-2
 ; Sequence 2, Application US/08923992A
 ; Patent No. 6280738
 ; GENERAL INFORMATION:
 ; APPLICANT: Tai, Joseph Y.
 ; APPLICANT: Blake, Milan S.
 ; TITLE OF INVENTION: No. 6280738-IGA Fc Binding forms of the Group B
 ; TITLE OF INVENTION: Streptococcal Beta Antigens
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/923,992A
 FILING DATE: 05-SEP-1997
 CLASSIFICATION: 536
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: US 60/024,707
 FILING DATE: 06-SEP-1996
 ATTORNEY/AGENT INFORMATION:
 NAME: Esmond, Robert W.
 REGISTRATION NUMBER: 32,893
 REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 371-2500
 TELEFAX: (202) 371-2540
 INFORMATION FOR SEQ ID NO: 2:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1164 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: protein
 US-08-923-992A-2

Query Match 7.2%; Score 172.5; DB 3; Length 1164;
 Best Local Similarity 21.6%; Pred. No. 0.0002;
 Matches 124; Conservative 77; Mismatches 189; Indels 185; Gaps 25;

QY 36 LSNERPLNE-----QTAEEDKIKTYTPENKPGQSNYSFVDNLLKAIT-EKEK 86
 Db 149 LELENQFNETNRLHLIKQHEVEVEKKAQ-----QKTLKQSDTKVDLSNIDKLNHQSQ 204
 QY 87 IEK--ERQSTIRSSPLDNKL-----NVEDVSTKNRKLIDDYDSTKSLDHK 130

Db 205 VERMAQKGTINDEKDSMLKKTIEDIRKQQAQDKKEDAEVKEEGLKLFSSSKAGLDQE 264
 QY 131 FQDDPDLHQLDGTPLTAEDIVHKI-----AARIYENDRAVDFDKIVSKLLNLG 179
 Db 265 IQE-----HVKET--SSBENTQKVDEHYANSIQNLAKSLEELDKATTNEQATQVKNQF 317
 QY 180 L-----ITESAHTLEDEVAEV-----LQKLISKE---ANNYEE 210
 Db 318 LENAQKLEIQPLIKETNVKLYKAMESLEQVEKELKHNSEANLEDLVAKSKEIVREYEG 377
 QY 211 DPNKPTSWTE-----NOAGKIPKVTWMAAIOQDGLAKGENDETVPN 351
 Db 378 KLNQSKNLDELQLEBEAHSKLVQVEDFRKFKTSEQVTPKRVKRDAAANENQ---Q 434
 QY 252 TLTTLNGLERRTKTYSSEDFELQVFPNFYALLKSIDSEKAEKETLITIMKTLIDFVK 311
 Db 435 KIELTVSPENITVYEGED-----VKFTVTAKS-DS-----KTTLDFSD 471
 QY 312 MMVKYG-TISPREGVSYLENLD-----EMIALQTK-----NLEKNATD 349
 Db 472 LITKYNPSVSDRISTNYKTNTDNHKAETITKNLKNESQVTLKAKODSGNVVEKTFPI 531
 QY 350 NISKLPAPSEKSHETDSTKEEAA-----KMEKEYGSLKD 385
 Db 532 TVQKKEEKQVKTPEQKDSKTEEKVPQBPKNQKQLELIKSAQOELEKLEKAIKELME 591
 QY 386 STKDONSNGGKTDPEKGTAYEALRAKNIWELK-----HDKGNKEDYDLSKMRDFIN 441
 Db 592 Q-PEIPSNP--EYGIKSIWESQKEPIQEAITSFKKIIGDSSSKYTYTEHYFNKYKSDPMN 648
 QY 442 KOADAVVEKGILDKKEAE-----AIKRIYSS 467
 Db 649 YOLHAQME--MLTRKVVQYMKNPDAEIKKIFES 681

RESULT 3

US-08-923-992A-8
 ; Sequence 8, Application US/08923992A
 ; Patent No. 6280738
 ; GENERAL INFORMATION:
 ; APPLICANT: Tai, Joseph Y.
 ; APPLICANT: Blake, Milan S.
 ; TITLE OF INVENTION: No. 6280738-IGA Fc Binding Forms of the Group B
 ; TITLE OF INVENTION: Streptococcal Beta Antigens
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/923,992A
 ; FILING DATE: 05-SEP-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/024,707
 ; FILING DATE: 06-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 8:


```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

Query Match
  7.0%; Score 169; DB 3; Length 1098;
Best Local Similarity 19.3%; Pred. No. 0.00034;
Matches 104; Conservative 80; Mismatches 182; Indels 174; Gaps 19;

Qy 25 GSODKSLHNRELSAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSF---VDNLNLKAI 81
Db 74 GKREKQLQWKNLKNVDNHTLSHEQKNBFKTKIDETNDSALLELENFNETNRLHI 133

Qy 82 TEKEKIEKERSRSP---DNKLNVEDVDSTKNR-----KLID-- 119
Db 134 KOHEVEKDKKAKQOKTLKQSDTKVDLSNIDKELNHQKSQAGITNEDKDSMLKIEDIR 193

Qy 120 -----YDSTKSGLDHFKQDDPGDGLHQLDGTPLTAEDIVHKI- 155
Db 194 KQAQOPDKKDAEVKVRBELGKLSSTKAGLDQEIQ-----HVKET--SSEENTQKVD 246

Qy 156 -----AARIYENDRAVDFKIVSKLNLGL-----ITESQA 186
Db 247 EHYANSLQNLAKSLEELDKAATNEQATQVKNQFLENAQKLEIQLIKETNVKLYKAMS 306

Qy 187 HTLEDEVAEV-----LQKLISKE--ANNYEEDPNKPTSWTE----- 220
Db 307 ESLEQVEKELKHNSFANLELDVAKSKEIVRYEGKLNQSKNLPKQLKEEAHAKLKQV 366

Qy 221 ---NQAGKIPKVTMAAIOQGLAKGENDETVSNLTITLNGLERRTKTSYEDNFEELYF 277
Db 367 EDPRKFKTSEQVTPKPKRLKDLAANNQ---OKIELTVSPENITVVEGED----- 415

Qy 278 PNFVALLKSIDSEKAKEKELITIMTKLIDPFVQWVKYV-TISPEEGVSVLENLD---- 332
Db 416 VKFTVTAKS-DS-----KTTLDFSDLLTKYNFVSDRISTNTKNTDNHKI 460

Qy 333 -----EMIALQTK-----NKLEKNATDNISKLPAPSEKSHSEEDSTKERA 375
Db 461 ABITIKNLKNESQVTLAKADSGNVVEKFTTITVQKKEKQVPTPEQKDSKTEKVP 520

Qy 376 MEKEYGSLKSTKDDNSPGGKTDPEKGTAYLEAIRKNIEWLKHKDKGNKEDYDLSK 435
Db 521 QEPK-----SNDKNQLOLEIKSAQOELEKLEKAIKELMEQPEIFSNPEYGIQK 568

RESULT 4
US-08-923-992A-10
; Sequence 10, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923.992A
; FILING DATE: 05-SEP-1997
```

```
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO. 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-10

Query Match
  7.0%; Score 168; DB 3; Length 1164;
Best Local Similarity 19.0%; Pred. No. 0.00044;
Matches 121; Conservative 85; Mismatches 177; Indels 254; Gaps 25;

Qy 25 GSQDKSLHNRELSAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSF---VDNLNLKAI 81
Db 105 GKREKQLQWKNLKNVDNHTLSHEQKNBFKTKIDETNDSALLELENFNETNRLHI 164

Qy 82 TEKEKIEKERSRSP---DNKLNVEDVDSTKNR-----K 115
Db 165 KOHEVEKDKKAKQOKTLKQSDTKVDLSNIDKELNHQKSQVETWAEQGITNEDKDSMLK 224

Qy 116 LIDD-----YDSTKSGLDHFKQDDPGDGLHQLDGTPLTAED 150
Db 225 KIEDIRKQAQADKEDAEVAVREELGKLTSTKAGLDQEIQ-----HVKET--TSEE 277

Qy 151 IVHKI-----AARIYENDRAVDFKIVSKLNL-----LGLITE---- 183
Db 278 NTQKVDEHYPNLSLQNLAKSLEELDKAATNEQATQVKNQFLENAQKLEIQLIKETNVK 337

Qy 184 -----SOAHTLEDEVAEVQLKLSKE--ANNYEEDPNKPTSWTE--- 220
Db 338 LYKAMESLEQVEKQLKHNSQAN--LEDLVAK-----SKEIVRYEGKLNQSKNLPKQ 390

Qy 221 -----NQAGKIPKVTMAAIOQGLAKGEND-----TVS--NTLTITNG 258
Db 391 LEEBAHSLKQVVEDPRKFKTSEQVTPKRVKEDLAANNQOKIELTVSPENITVVEG 450

Qy 259 LERTKTSYEDNFEELYFNFVALLKSIDSEKAKEKELITIMTKLIDPFVQWVKYV- 317
Db 451 -----EDLK-----FTLTAKSDS-----KTTLDFSDLLTKYNP 478

Qy 318 TISPEEGVSVLENLD-----EMIALQTK-----NKLEKNATDNISKLP 356
Db 479 VSDRISTNTKNTDNHKAIBITIKNLKNESQVTLAKADSGNVVQKFTTITVQKKE 538

Qy 357 APSEKSHSEEDSTKEEAAKMEKEYGSLKSTKDDNSPGGKTDPEKGTAYLEAIRKNI 416
Db 539 KOVPKTPQKDSKTEKVPQEPK-----SNDKNQLOLEIKSAQOELEKLEKAI 586

Qy 417 EWLKHKDKGNKEDYDLSK-----MRDF 439
Db 587 KELMEQPEIFSNPEYGIQKSQWESQKEPIQBAITSFKKIIGDSSSKYVTEHYFNKYKSD 646

Qy 440 INKQADAYVEKGILDKEAE-----AIKRIYSS 467
Db 647 MNYQLHAQME--MLTRKVQVQYINKYPADNAEIKKIFES 681

RESULT 5
US-09-914-259-16
; Sequence 16, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
```

```

; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 957
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-914-259-16

Query Match      7.0%; Score 167.5; DB 4; Length 957;
Best Local Similarity 20.3%; Pred. No. 0.00036;
Matches 103; Conservative 98; Mismatches 182; Indels 125; Gaps 23;

QY 25 GSQDKSLNR-----ELSAERPLNEQIAEAEEDKIKTYPENKPGQSNYSFVDNLNL- 78
Db 321 QORAKTIKNTSVNLELTAE-----EMKKYKEKEKNTLKNVIOHLEML 367
QY 79 -----KATKEKEKIERQIRSPDLNKLNVEDVD-----STNRKLIDYDSTKSG 126
Db 368 NRWRNGEAVPEDEQISAKDQK-NLEPCDNTPIIDNIAPVAVAGISTEEK--EKYDEISS 423
QY 127 LDHKFQDDPDGLHQLDGTPLTAEDIVHKIAIRIYEEND-----RAVFDKIVSKLLNLGLI 181
Db 424 LYRQLDDKDRINO-----QSQAELKIQQMDQDELLASTRDYKEIQEELTRQIE 476
QY 182 TESQAHTLEDVAEVLQKLISKEANNYE---EDPNKPTSWTNOAGKIPKVPMAAIQD 238
Db 477 NEA-----AKDEVEVLOAL-EELAVNYDQKSQEVEDKTRANEQLTDELAQKTTLTITQR 531
QY 239 GLAK-----GENDETVSNTLITNGLERR-----TK 264
Db 532 ELSQQLSLSNHQKRATEINLLKDLGEIGGIITND---VKTADVNGVIEEFTMAR 588
QY 265 TYSEDNFEELQYFNFYALLKS--IDSEKAEKETLITIMTKLIDFVKMVKYGVISPE 322
Db 589 LYISKMKSEVKSLSVNRSKQLESQMSDNKKNKASERELACQLLSQHEAKIKSLT----- 644
QY 323 EGVSYLENDEMIALQYFNKLEK---NATDNISKLFPAPSEKSHETDSTKE----- 371
Db 645 ---DYMOMME-----QKRRLQESQDSLSBELAKL--RAQEKHVEVSFQDKEKHLRLQ 694
QY 372 EAAKMEKEYGSLKDKSTDDNSNPGKTDPEPKTEAVLEAIRKNIEWLKHDKKGNKEDY 431
Db 695 DAEEMKALQWESHEAHQKLSRLRDEIEBKQIIDEIRLNQKLQLEQEK-LSSDY 753
QY 432 DLSQMRDFINKQADAVVEKGIK--DKEE 457
Db 754 NKLKIED---QEREMKLEKLLNDKRE 778

```

```

RESULT 6
US-08-923-992A-4
; Sequence 4, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IGA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESS: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Emond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1104 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-4

Query Match      6.9%; Score 166; DB 3; Length 1104;
Best Local Similarity 21.2%; Pred. No. 0.00057;
Matches 121; Conservative 80; Mismatches 190; Indels 181; Gaps 26;

QY 17 IQAFPKPGSQDSKSLHNRRL-----SARPLNEQIAEAEEDKIKTYPENKPGQSNYSF 71
Db 139 VEKHNKP--NQOKTLKQSDTKVDLSNIDKELNHQKSQVEA-----MAEQAGITN 185
QY 72 VDNLLKKAITEKEKIKERQIRSPDLNKLNVEDVDSTKNRKLIDYDSTKSLDGHKF 131
Db 186 EDKDSMLKKI---EDIRKQAQ-----ADKK---EAEVKVREELGKLFSSTKAGLDQOI 234
QY 132 QDDPGLHQLDGTPLTAEDIVHKI-----AARIYENDRAVFDKIVSKLLNLGL 180
Db 235 QE-----HVKKET--SSEENTQKVDEHYANSLQNLAKSLEELDKATTNEQATQVKNQFL 287
QY 181 -----ITESQAHTLEDEVAEV-----LOKLISKE-----ANNYEED 211
Db 288 ENAQKLEKIQLIKETNVKLYKAMSESLEQVEKELKHNSEANLQDLVAKSEIIVREYEGK 347
QY 212 PNKPTSWTE-----NQAGKIPKVPMAAIQDGLAKGENDETVSNT 252
Db 348 LNQSKNLPQLBEEAHSKLVQVVEHPRKFKTSEQVTPKRVKRDLLAANNQ---QK 404
QY 253 LTLTNGLERRTKYTSEDNFEELQYFNFYALLKSIDSEKAEKETLITIMTKLIDFVKM 312
Db 405 IELTVSPENITVYEGED-----VKPTVTAKS-DS-----KTTLDPSDL 441
QY 313 MVKYG-TISPEEGSYLENLD-----EMIALQTK-----NKLEKNATDN 350
Db 442 LTKYNPSVSDRISYNTKINTDNHKAETIKNKLKNSQTVTLAKADDSGNVVEKTFIT 501
QY 351 I-----SKLFPAPSEKSHETD-----STKEAAKMEKEYGSLKDS 386
Db 502 VQKKEEKQVPTPEQKHSKTEQNVQPSKNDKNQQLBELIKSAQQLKLEKAIKELMEQ 561
QY 387 TKDQNSNP--CGKTDPEKGTAEVLAIRKNIEWLKHDKKGNKEDYDLSQMRDFINKQA 444
Db 562 -PEIPSNPEYGIQKSIWESQKEPIQEAITSFNKTIIGSSSSKITYTEHYFNKYKSHFMNYQL 620
QY 445 DAYVEKGILDKBEAE-----AIKRIYSS 467
Db 621 HAQME--ILTRKVVQYMNKYPDNAEIKKIPES 650

```

RESULT 7

```

US-08-923-992A-6
; Sequence 6, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923.992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 6:
; LENGTH: 1128 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-6

Query Match 6.9%; Score 166; DB 3; Length 1128;
Best Local Similarity 19.3%; Pred. No. 0.00059;
Matches 119; Conservative 86; Mismatches 200; Indels 210; Gaps 23;

Qy 25 GSQKSLNRLSARPLNEQIAEEDKIKKTPPENKPGQSNVSP---VDNLLKAI 81
Db 69 GKREKQLQWKNLKNVDNTILSHEQKNEFKTKIDETNDSALLELENQFNETNRLHI 128
Qy 82 TEKEKIEKERSIRSP---DNKLVNEDVDSTNR-----K 115
Db 129 KQBEVEKDKKAKQOKTLKQSDTKVDLSNIDKELNHQKSPVEKMAEPKGINEDKSNLK 188
Qy 116 LIDD-----YDSTKSGLDHFKQDDPDGDLHOLDGTPLEAED 150
Db 189 KIEDIRKQAQADKEDAEVKBRELKGLFSTKAGLDQE-----IHEVVKETSSE 241
Qy 151 IVHKI-----AARIYEENDRAVDKIVSKLNLGL----- 180
Db 242 NTQKVDEHYANSLQNAQKLEELDATTNEQATQVKNQFLENAQKLKEMQPLIKETNVK 301
Qy 181 ITESQAHTELEDAEV-----LQKLISKE---ANNYEEDPNKPTSWTE----- 220
Db 302 LYKAMSLEQVEKELKHNSANLELDLVAKSKEIVREYEGKLNQSKNLPKQLKEEBAHS 361
Qy 221 -----NOAGKPIPEKVTPEMAAIQGLAKGENDETSTNTLTNGLERRTKTSYDNF 271
Db 362 KKKQVVEDFRKKFTSEQVTPKVKVRDLAENNQ---QKIELTVSPENITVVEGED-- 416
Qy 272 EELQYFNFYALLKSIDSEKAKEKETLITIMKTLIDFVKQWVKYG-TISPEGVSYLEN 330

```

```

Db 417 -----VKFTVTAKS-DS-----KTTLDPSDLLTKYNPSVSDRISTNYKTN 455
Qy 331 LD-----EMIALQTK-----NKLEKNATONISKLFPPAPSEKSHEDTST 369
Db 456 TDNHKIAEITIKNLKINESQVTLKAKDSDGNVVVEKFTTITVQKKEKQVPEKTEQKDSK 515
Qy 370 KEAA-----KMEKEYGSLKOSTKDDNNDNPGKGTDEPKGT 405
Db 516 TEKVQPEKPSNDKNQQLQELIKSAQOELEKAIKELMEQ-PEIPSNP--EYGTQKSIW 572
Qy 406 EAYLEATRKNLEWLK-----HDKKNGKEDYDLSKWRDFINKQADAYVEKGLDKKEAB-- 459
Db 573 ESQKEPIQEAITSFPKIIGDSSSKYTYTEHYFNKYKSDFMNYQLHAQME--MLTRKVVQYM 630
Qy 460 -----AIKRIYSS 467
Db 631 NKYPDNAEIKKIPES 645

RESULT 8
US-09-595-684B-31
; Sequence 31, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cvt036
; CURRENT APPLICATION NUMBER: US/09/595.684B
; CURRENT FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 31
; LENGTH: 2662
; TYPE: PRT
; ORGANISM: Human
US-09-595-684B-31

```

```

Query Match 6.9%; Score 166; DB 4; Length 2662;
Best Local Similarity 21.8%; Pred. No. 0.0019;
Matches 119; Conservative 103; Mismatches 175; Indels 150; Gaps 28;

Qy 26 SQKSLNRLSARPLN---EQIAEAEEDKIKKTPPENKPGQSNYSFVDNLLKAIT 82
Db 1101 AQEK---NHAIKKEGELSRTCRLAEVVEEKLKESQQLQKQQL-----LVQEMS 1150
Qy 83 EKKEKIEKERSIRSPDLNKLNVEDVDSTK---NRKLIDDYDSTKS-----GLDHK 130
Db 1151 EMOKKINEIENKLNKELTLEHMETERLELAQKLNENYEVKSTKERRKVLKELQKS 1210
Qy 131 FQDDPDGL---HQLDGTPL-TAEDI-----VH-KIARIYEENDRAVDKIVSKLNLGL 180
Db 1211 FETERDHLRGYIREIEATGLQTKELKIAHILKHOETIDELRRSVSEK-TAQIINTQD 1269
Qy 181 ITESQAHTELEDAEVL---QKLIS--KEANNYEEDPNKPTSWTENQAGKPIPEKVTPEMAA 235
Db 1270 LEKS---HTKLOEIRPVLHEEQELLPNVKYSETQETNNELELITEQSTTK---DSTTLAR 1324
Qy 236 IQ-----DGLAK-GENDETSTNTL-----TLTNGLERRTK 264
Db 1325-1EMERLRNLEKFOESQBEIKSLTKERDNLKTIKEALEVKHDQLKEHIRETLAKIOESQSK 1384
Qy 265 TYSEDNFEEL-----QYFPNFYALLK-----SIDSEK-AKEKE 297
Db 1385 QEOSLNKKEKNETTKIVSEMEQFKPKDSALLRIEIMGLSKRLQSHDEMKSVAKEKD 1444

```

QY 298 TLITIMKTLIDFVKMMVKYGTISPEGVSYIENLDEMIA--LQTKNKLK-----KNATD 349
 Db 1445 DLORLQEVL-----QSEDLQAKENIKIIVAKHLETEBELKVAHCCKLKEOE 1490
 QY 350 NISLFPAPSEKSHETDSTKEEAKKEKEYGSLKOSTKDNPNPGKTDPEPKTE--- 406
 Db 1491 TINELRVNLSEK-----STEISTIOKQLEAINDKLON-----KIOEYKEBEQLN 1535
 QY 407 -AYLEAIRKNIEWLKKHDKGNKEDYDL-----SKMROFINKQADAYVEKGILDKKEAEAI 461
 Db 1536 IKQISEVQENVNELKQPKHEKAKDSALQSIESKMLELTNRLQESQBEIQIMIKEX-EM 1594
 QY 462 KRIYSSL 468
 Db 1595 KRVQEAL 1601

RESULT 9
 US-09-914-259-17
 ; Sequence 17, Application US/09914259
 ; Patent No. 6495336
 ; GENERAL INFORMATION:
 ; APPLICANT: Makowski, Lee
 ; APPLICANT: Hyman, Paul
 ; APPLICANT: Williams, Mark
 ; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
 ; FILE REFERENCE: 8471-010-999
 ; CURRENT APPLICATION NUMBER: US/09/914,259
 ; CURRENT FILING DATE: 2000-11-21
 ; NUMBER OF SEQ ID NOS: 180
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 17
 ; LENGTH: 956
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 US-09-914-259-17

Query Match 6.8%; Score 163.5; DB 4; Length 956;
 Best Local Similarity 20.7%; Pred. No. 0.00072;
 Matches 106; Conservative 91; Mismatches 169; Indels 147; Gaps 25;
 QY 25 GSQDQSLHNR-----ELSAERPLNEQIAEAEEDKIKKTYPPENKPGQSNYSFVDNLLNL- 78
 Db 321 GQAKTIKTVSNVLELTAB-----EWKKYKEKEKKNALKSVLQHLEML 367
 QY 79 -----KAITEKIE-KERQSISSPLDNKLNVED-----VDSTKRNKLIDDYDSTKSGL 127
 Db 368 NRWRNGEAVPEDEQISAKDHKSL--EPCDNTPIIDNITPVWDGISAEX--EKYDEEITS 423
 QY 128 DHKFQDDPDGLHOLDGTPLTAEDIVHKAARIYEEND-----RAVFDKIVSKLLNLGLIT 182
 Db 424 YRLDDKDDINO-----QSQAELKLOOMLDQDELLASTRRDYEKIQEELTRLOIEN 476
 QY 183 ESQAHTLEDEVAEVLQKLISKEANNYE--EDPNKPTSWENQAGKIPKVTFMAIQQG 239
 Db 477 EA-----AKQEVKEVLQAL--BELAVNDQKSQVEDKTRANEQLTDLAQKTLTTITTTQRE 531
 QY 240 LAK-----GENDETVSNTLITNLGLERR-----TKT 265
 Db 532 LSQLOELSNHQRKATEILNLLKOLGEIGGIITND---VKTLDVNGVIEEFETMARL 588
 QY 266 YSDNFEELQYFNFYALLKS--IDSEKAKEKETLITIMTKLIDFVKMMVKYGTISPEE 323
 Db 589 YISKMKSEKSLVNRKQLESQMDNRRKNNAASERELAAQCOLLSQHEAKIKSLT----- 643
 QY 324 GVSYLENDEMIALQTKNLEK---NATDNISKLPAPSEKSHET--DSTKEEAAKMEK 378
 Db 644 ---DYMQME-----QKRRLQESQDSLSBELAKL--RAQEKHVSFQDKEKHLTRLO- 693
 QY 379 EYGLSKDSTKDDNSNPGGKTDPEKGTETAYLBAIRKNIEWLKKHDKGNKEDYDLSKMRD 438
 Db 694 -----DAEEVKALEQQMESHR-----BAHQKQ-----LSRLRD 722

QY 439 FINKQADAYVEKGILDK-----EAEAIKRIYSSL 468
 Db 723 EIEEKQRIIDIRDNQKLEQLQERLSSDYNKL 755
 RESULT 10
 US-08-714-741-32
 ; Sequence 32, Application US/08714741
 ; Patent No. 6500613
 ; GENERAL INFORMATION:
 ; APPLICANT: Briles, David E.
 ; APPLICANT: McDaniel, Larry S.
 ; APPLICANT: Swiatlo, Edwin
 ; APPLICANT: Yother, Janet
 ; APPLICANT: Crain, Marilyn J.
 ; APPLICANT: Hollingshead, Susan
 ; APPLICANT: Tart, Rebecca
 ; APPLICANT: Brooks-Walter, Alexis
 ; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
 ; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
 ; TITLE OF INVENTION: PORTIONS AND PRODUCTS
 ; NUMBER OF SEQUENCES: 47
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Curtis, Morris & Safford, P.C.
 ; STREET: 530 Fifth Avenue
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA: US/08/714,741
 ; FILING DATE: 16-SEP-1996
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Frommer Esq., William S.
 ; REGISTRATION NUMBER: 25,506
 ; REFERENCE/DOCKET NUMBER: 454312-2460
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 840-3333
 ; TELEFAX: (212) 840-0712
 ; INFORMATION FOR SEQ ID NO: 32:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 8991 amino acids
 ; TYPE: amino acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: amino acid
 US-08-714-741-32

Query Match 6.8%; Score 162; DB 4; Length 8991;
 Best Local Similarity 18.2%; Pred. No. 0.02;
 Matches 88; Conservative 110; Mismatches 184; Indels 102; Gaps 19;
 QY 26 SQDKSLHNRLELSARPLNEQIAEAEEDKIKKTYPPENKPGQSNYSFVDNLLNLKAITEKE 85
 Db 2024 AQEKYADYQRIEBAKEAQESLEQAEANKDYOLKLUKYLDDGRNLNSSLVKKEMEAE 2083
 QY 86 KIEKERSQIRSSPLDNKL-----NVEDVDSTKRNKLIDDYDSTKSGLDHFKFD----- 133
 Db 2084 KKDKEKQA-----EFNKIRREIVVPNPQLELMARRKSEV--VKAKESGLVKRVEAEKV 2136
 QY 134 -----DPDGLHOLDGTPLTAEDIVHKAARIYEENDRAVFDKIVS-KLLNLGLITESQ 185
 Db 2137 TEARQKLDAAEAKEVLPQTRVENEVHKLQKMNKK--KMILTSLASVAILGAGLVTSQP 2194
 QY 186 AHTLEDEVAEVLQK-----LISKEANN-----YEEEDPNKPTSWT 219
 Db 2195 TFVRAEESPVQVEKSSLEKKYEEAKAKADTAKDYETAKKAEDAQKYEEDQKR-----T 2250

QY 220 ENQAGKIPE---KVTPTMA-----AIQDLAKGENDETVSNTLT-LTNGLERRTK 264
 Db 2251 EEKARKEAASQKLDIVLVQVNAVYKBYREVQNRQSKSDADYQKKLTETVDSKIEKARK 2310
 QY 265 TYS--ENFBELOVFPNFYALL---KSIDSEKAEKETLITIMTKLIDFVQMMVKYGT 318
 Db 2311 EOODLQNNFVR-----AVVAPDPTCVGDXDMNKKMILTSLSV-----AILGAGX 2359
 QY 319 ISPEEGVSVLENDEMIALQTKNLEKATNISKLPAPSEKSHETSTKKEAAKWEK 378
 Db 2360 VTSOPTXVRAEAPQV---EKSLEKKYBEAKV-----DAAKDYDEAKKKAQAQK 2411
 QY 379 EYSLKSTXDDNSNPGKTDPEKGTAEVLEAIRKNIEWKHDKKGN---KEDYDLK 435
 Db 2412 KYEDQKKTBEKAKAKAASEIAKATEEVQKAVLDYITAIRNHDGKTSABEAENKAK 2471
 QY 436 MRDF 439
 Db 2472 ERDY 2475

RESULT 11
 US-08-533-306A-6

; Sequence 6, Application US/08533306A

; Patent No. 5837457

; GENERAL INFORMATION:

; APPLICANT: Liu, Pu

; APPLICANT: Collins, Francis S.

; APPLICANT: Siciliano, Michael J.

; APPLICANT: Claxton, David

; TITLE OF INVENTION: Markers for Detection of Chromosome 16

; TITLE OF INVENTION: Rearrangements

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

; STREET: P.O. Box 828

; CITY: Bloomfield Hills

; STATE: MI

; COUNTRY: USA

; ZIP: 48303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/533.306A

; FILING DATE: September 25, 1995

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, Deann F.

; REGISTRATION NUMBER: 36683

; REFERENCE/DOCKET NUMBER: 2115-00869COB

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (810) 641-1600

; TELEFAX: (810) 641-0270

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 816 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; US-08-533-306A-6

Query Match

Best Local Similarity 22.6%; Pred. No. 0.0011; Length 816;

Matches 109; Conservative 81; Mismatches 193; Indels 99; Gaps 21;

QY 26 SQKSLHNRRLSRRPLNEQIARBEKIKKTPPENKPGQSNYSFV--DNMLLKAITE 83

Db 158 SHREEMENEVSVTGMLNE--AEGKATKLAKDVASLSQLDQTQLIQQEETRQKLVSTK 215

QY 84 KEKIEKROSIR---SSPLDNKLNVEDVDSTKN-----RKLIDDYDSTKSGLDH---KF 131
 Db 216 LRQLEERNLQDLQDEMEAKQNLERHISTLNIQLSDSKKQLQDFASTVEALEEGKRF 275
 QY 132 QDDPDGLHQLDGTPLTAEDIVHKAIRIYENDRAVPD-----KIVSKLNLGLITESQA 186
 Db 276 QKSIENLTQOYEKAAYDKLETKNRLQQLDLDLVDLDNQQLVSNL-----EKQ 328
 QY 187 HTLEDEVAEVLQKLISKEANNYBEDPNKPTSTWENQAGKIPEKVTPTMA-AIQDLAKGEN 245
 Db 329 RKFDQLLAE--EKNIS---SKYADERDR---AEEAREKETKALSARALEEALAEKEE 379
 QY 246 DETVSNTLTITNGLERTKTYSERD-----NFEELQYFPNFYALLKSIDSEKAEKETLIT 301
 Db 380 LERTNMLK----AEMEDLVSSKDDVGKNVHEL-----EKSRALETQME 420
 QY 302 IMKTLIDFVQMMVKYGTISPEEGVSVLENDEMIALQTKNLEKATNISK-----L 354
 Db 421 EMKTLQEELEDELQ-----ASEDAKLRLVNMQALKGQFERDLQARDQNEEKRLQRL 476
 QY 355 FPAPSEKSHETDSTKEEAAKMEKEYGSLKXDSKTDNSNPGKTDEPK--GKTEAYLEAI 412
 Db 477 HEYTELEDEDERNERALAAAKKLE-GDLKDLQADSIAKGRRAIKQLRKLOAQMKDF 535
 QY 413 RKNIE-----WLKHKDKGNKEDYDLKMRDFI-----NKQADAYVEKGLDK 455
 Db 536 QRELEDARASRDEIFATAKENEKAKSLEADLMQLQEDLAAAEARAKQAD-----LEK 588
 QY 456 EE 457
 Db 589 EE 590

RESULT 12

US-08-742-923A-6

; Sequence 6, Application US/08742923A

; Patent No. 5869611

; GENERAL INFORMATION:

; APPLICANT: Liu, Pu

; APPLICANT: Collins, Francis S.

; APPLICANT: Siciliano, Michael J.

; APPLICANT: Claxton, David

; TITLE OF INVENTION: Markers for Detection of Chromosome 16

; TITLE OF INVENTION: Rearrangements

; NUMBER OF SEQUENCES: 14

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

; STREET: P.O. Box 828

; CITY: Bloomfield Hills

; STATE: MI

; COUNTRY: USA

; ZIP: 48303

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: PatentIn Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/742,923A

; FILING DATE: No. 5869611ember 1, 1996

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: Smith, Deann F.

; REGISTRATION NUMBER: 36683

; REFERENCE/DOCKET NUMBER: 2115-00869DVC

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (810) 641-1600

; TELEFAX: (810) 641-0270

; INFORMATION FOR SEQ ID NO: 6:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 816 amino acids

; TYPE: amino acid

; TOPOLOGY: linear

```

; MOLECULE TYPE: protein
; US-08-742-923A-6
;
; Query Match      6.6%; Score 159.5; DB 2; Length 816;
; Best Local Similarity 22.6%; Pred. No. 0.0011;
; Matches 109; Conservative 81; Mismatches 193; Indels 99; Gaps 21;
;
; 26 SQDKSLHRELNSAERPLNEQIAEAEEDKIKKTYPPENKPGQSNYSFV--DNLNLLKAITTE 83
; 158 SHREEMENEVESVTGMLNE--AEGKAIKLAKDVASLSQLQDTQELLQEEETROKLNVTSTK 215
; 84 KEKTEKERQSTR---SSPLDNKLNVEDVDSTKN-----RKLIDDYDSTKSGLDH---KF 131
; 216 LRQLEERNSLQDQDDEMEAKQNLERHISTLNQLSDSKKKLQDFASTVEALEEGKKRF 275
; 132 QDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFD-----KIVSKLLNLGLITESQA 186
; 276 QKEIENLTQQYEEKAAAYDKLEKTKNRLOQELDLVVDLNDQRLVSNL-----EKKQ 328
; 187 HTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTTPMA-AIQDGLAKGEN 245
; 329 RKFDQLLAE--EKNIS---SKYADERDR---AAEAAREKETKALSARALEEAEKEE 379
; 246 DETVSNLTNLGLERRTKTYSED---NFEELOYPFNFYALLKSIDSEKAKEKETLIT 301
; 380 LERTNKMKL---AEMEDLVSSKDDVGKNVHEL-----EKKRALETQME 420
; 302 IMKTLIDPVKMMVKYGTISPREGVSYLENDEMIALQTKNLEKNATDNISK-----L 354
; 421 EMKTLQEELEDELQ---ASEDAKLRLEVNMQALKGQFERDLQARDEQNEKKRQLQRL 476
; 355 FPAFSEKSHETDSTKEBAAKMEKEYGSLKDDSTKDDNSNPGGKTDEPK--GKTEAYLEAI 412
; 477 HEYTELEDERNERALAAAKKLE-GDLKDLQADSAIKGREBAIKQLRKLQAKMDF 535
; 413 RKNIE-----WLKXHDKGNKEDYDLSKMRDFI-----NKQADAYVEKGILDK 455
; 536 QRELEDAARSDEIPATAKENEKKAKEADLMQLQEDLAAAEARAKQAD-----LEK 588
;
; 456 EE 457
; 589 EE 590
;
; RESULT 13
; US-08-533-306A-4
; Sequence 4, Application US/08533306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533.306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
;
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869COB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-533-306A-4
;
; Query Match      6.6%; Score 158.5; DB 2; Length 885;
; Best Local Similarity 22.9%; Pred. No. 0.0015;
; Matches 109; Conservative 79; Mismatches 190; Indels 99; Gaps 21;
;
; 31 LHNRELSAERPLNEQIAEAEEDKIKKTYPPENKPGQSNYSFV--DNLNLLKAITTEKIE 88
; 232 LQNEVESVTGMLNE--AEGKAIKLAKDVASLSQLQDTQELLQEEETROKLNVTSTKRLQ 289
; 89 KERQIR---SSPLDNKLNVEDVDSTKN-----RKLIDDYDSTKSGLDH---KFQDDPD 136
; 290 BERNLQDQDDEMEAKQNLERHISTLNQLSDSKKKLQDFASTVEALEEGKKRFQKEIE 349
; 137 GLHLDGTPLTAEDIVHKIAARIYEENDRAVFD-----KIVSKLLNLGLITESQAHTLED 191
; 350 NLTQQYEEKAAAYDKLEKTKNRLOQELDLVVDLNDQRLVSNL-----EKKQKRFQ 402
; 192 EVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTTPMA-AIQDGLAKGENDETVS 250
; 403 LLAE--EKNIS---SKYADERDR---AAEAAREKETKALSARALEEAEKEELERTN 453
; 251 NTLTITNLGLERRTKTYSED---NFEELOYPFNFYALLKSIDSEKAKEKETLITIMKTL 306
; 454 KMLK---AEMEDLVSSKDDVGKNVHEL-----EKKRALETQMEEMKTO 494
; 307 IDFKMMVKYGTISPREGVSYLENDEMIALQTKNLEKNATDNISK-----LFPAPS 359
; 495 LEELEDELQ---ASEDAKLRLEVNMQALKGQFERDLQARDEQNEKKRQLQRLHVEYET 550
; 360 EKSHEETDSTKEBAAKMEKEYGSLKDDSTKDDNSNPGGKTDEPK--GKTEAYLEAIRKNE 417
; 551 ELEDERNERALAAAKKLE-GDLKDLQADSAIKGREBAIKQLRKLQAKMDFQRELE 609
; 418 -----WLKXHDKGNKEDYDLSKMRDFI-----NKQADAYVEKGILDKEE 457
; 610 DARASRDEIFATAKENEKKAKEADLMQLQEDLAAAEARAKQAD-----LEKEE 659
;
; RESULT 14
; US-08-742-923A-4
; Sequence 4, Application US/08742923A
; Patent No. 5869611
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

```

```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742,923A
; FILING DATE: No. 5869611 member 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-742-923A-4

Query Match 6.6%; Score 158.5; DB 2; Length 885;
Best Local Similarity 22.9%; Pred. No. 0.0015;
Matches 109; Conservative 79; Mismatches 190; Indels 99; Gaps 21;

Qy 31 LHNRELSAERPLNQLAEAEEDKIKKTPPENKPGQSNYSFV--DNLLKKAITEKEKIE 88
Db 232 LQNEVESVTGMLNE--AEQKAIKLAKDVASLSQLODTQELLQEBETROKLNVNSTKLRLQE 289

Qy 89 KEROSIR--SSPLDNKLNVEDVSTKN-----RKLIDDYDSTKGLDH---KFODDPD 136
Db 290 EERNLQQLDEEMAKONLHRHSTLNLIQSDSKKLQDFASTVEALBEGKGRFQKEIE 349

Qy 137 GLHQLDGTPLTAEDIVHKIAARIYEENDRAVFD-----KIVSKLLNLGLITESQAHTLED 191
Db 350 NLTQQYEKAAYDKLETKNRLQELDDLVDLDNQQLVSNL-----EKKQKRFQ 402

Qy 192 EVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTMA--AIQDGLAKGENDETVS 250
Db 403 LLAE--EKVNIS--SKYADERDR---AAEAAREKETKALSARALBEALBEALEKELERTN 453

Qy 251 NTLTLTNGLERRTKTYSED---NFEELQYFPNFYALLKSIDSEKEKETLITIMKTL 306
Db 454 KVLK-----AEMEDLVSSKDDVGKVNHEL-----EKSRALETQMEEMKTQ 494

Qy 307 IDFKVMVKYGTISPESGVYLENLDLMIALQTKNLEKNATDNISK-----LFPAPS 359
Db 495 LEELEDELQ-----ASEDAKLELVNMWALKQGFERDLQARDEQNEEKRRQLORQHYET 550

Qy 360 EKSHEETDSTKEAAKMEKEYGSLKSDTKDSDNSPGKTDDEPK--GKTEAYLEAIRKNIE 417
Db 551 ELEDERNERAAAAAKKLE--GDLKDLQLADSALIKGREAAIKQLRLQAQMKDFORELE 609

Qy 418 -----WLKHKDKGKNKEDYDLSKMRDFI-----NQADAVVEKGILDKEE 457
Db 610 DARASRDEIFATAKENEKKAKEADLMQLQEDLAAAERARKQAD-----LEKEE 659

```

RESULT 15

```

US-09-914-259-12
; Sequence 12, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12

```

```

; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
US-09-914-259-12

Query Match 6.5%; Score 156.5; DB 4; Length 1087;
Best Local Similarity 19.3%; Pred. No. 0.0028;
Matches 103; Conservative 83; Mismatches 180; Indels 167; Gaps 22;

Qy 31 LHNRELSAERPLNQLAEAEEDKIKKTPPENKPGQSNYSFVNDLNLKKAITEKEKIEKE 90
Db 304 IHEKEVEIDR-LNEQIIKLOQQ-----LKITTDKNKVIEEK 337

Qy 91 QSISSPLDNKLNVEDVST--KNR-----KLIDDYDSTKGLDHKFODDPDGLHQL 141
Db 338 NELIRDLEAQIECLMSQOERVRKNREBEIEQLNEVIEKLOQELANIQKTSVDPSSLSL- 396

Qy 142 DGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVA----- 194
Db 397 -----EADSLKHQL-----DKVIAEKLAL-----EHQVETTNEMAVTKNVLK 434

Qy 195 -----EVLQKLIS--KEANNYEEDPNKPTSWTENQAGKIPEKVTMAAI--ODGLAKG 243
Db 435 ETNFKMQLTQELCSLKRERKMERIQSVPEKSVNMSVGLSKDKPKENDLIPTEDALAQL 494

Qy 244 E-----NDETVSNTL--TLTNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKE 295
Db 495 ETQTQLRSSESSKVSLSLETLLQLESTVSTKDLTLTCY-----KQIQDMRQGR 547

Qy 296 KET-----LITIMKTL-----IDFKVMVKYGTISPESG----- 324
Db 548 SETEMLTQTKIVSLQKVLKEEKAALVSQVLEAVQYVVKLCADKPAVSDPAPTEVPGLS 607

Qy 325 -----VSYLENL-----DEMIALQTKNLEK-NATDNISKLF 355
Db 608 QLAGNTMESDVSAITWRISELSQVEMHSSLISEKEQVIAEKNALEKKEKLOELQKLV 667

Qy 356 PAPSEKSHEETDSTKEAAKMEKEYGSLKSDTKDSDNSNPG-----GKTDEPKGKTBAYL 409
Db 668 QDSETKQRE-----RERQSRHLGDLGVLESTTSEESGVFGELEALRAESAAPKGE LANYK 722

Qy 410 EATRKNIEWLKKHDKGKNKEDYDLSKMRDFINQKQADA---YVEKGILDKEEAE 459
Db 723 ELAEKLOEELLVKETNMASLPKLSHVRLDQTEAEDKLSHFSEKE--DKTEVQ 773

Search completed: January 2, 2004, 12:36:20
Job time : 22 secs

```

THIS PAGE BLANK (USPTO)


```

; PRIOR APPLICATION NUMBER: 60/125774
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125778
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/125826
; PRIOR FILING DATE: 1999-03-24
; PRIOR APPLICATION NUMBER: 60/127035
; PRIOR FILING DATE: 1999-03-31
; PRIOR APPLICATION NUMBER: 60/127706
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/129122
; PRIOR FILING DATE: 1999-04-13
; PRIOR APPLICATION NUMBER: 60/130359
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131272
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/132371
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/132379
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/132383
; PRIOR FILING DATE: 1999-05-04
; PRIOR APPLICATION NUMBER: 60/135750
; PRIOR FILING DATE: 1999-05-25
; PRIOR APPLICATION NUMBER: 60/138166
; PRIOR FILING DATE: 1999-06-08
; PRIOR APPLICATION NUMBER: 60/144791
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/146970
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/162506
; PRIOR FILING DATE: 1999-10-29
; PRIOR APPLICATION NUMBER: 09/311832
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 09/380142
; PRIOR FILING DATE: 1999-08-25
; PRIOR APPLICATION NUMBER: 09/644848
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 09/747259
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: 09/816744
; PRIOR FILING DATE: 2001-03-22
; PRIOR APPLICATION NUMBER: 09/854208
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/854280
; PRIOR FILING DATE: 2001-05-10
; PRIOR APPLICATION NUMBER: 09/874503
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: 09/869599
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: 09/908,827
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: PCT/US99/10733
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: PCT/US99/28551
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30720
; PRIOR FILING DATE: 1999-12-22
; PRIOR APPLICATION NUMBER: PCT/US00/05601
; PRIOR FILING DATE: 2000-03-01
; PRIOR APPLICATION NUMBER: PCT/US00/05841
; PRIOR FILING DATE: 2000-03-02
; PRIOR APPLICATION NUMBER: PCT/US00/14042
; PRIOR FILING DATE: 2000-05-22
; PRIOR APPLICATION NUMBER: PCT/US00/15264
; PRIOR FILING DATE: 2000-06-02
; PRIOR APPLICATION NUMBER: PCT/US00/23522
; PRIOR FILING DATE: 2000-08-23
; PRIOR APPLICATION NUMBER: PCT/US00/23328

```

```

; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: PCT/US00/32678
; PRIOR FILING DATE: 2000-12-01
; PRIOR APPLICATION NUMBER: PCT/US00/34956
; PRIOR FILING DATE: 2000-12-20
; PRIOR APPLICATION NUMBER: PCT/US01/06520
; PRIOR FILING DATE: 2001-08-28
; PRIOR APPLICATION NUMBER: PCT/US01/17800
; PRIOR FILING DATE: 2001-06-01
; PRIOR APPLICATION NUMBER: PCT/US01/19692
; PRIOR FILING DATE: 2001-06-20
; PRIOR APPLICATION NUMBER: PCT/US01/21066
; PRIOR FILING DATE: 2001-06-29
; PRIOR APPLICATION NUMBER: PCT/US01/21735
; PRIOR FILING DATE: 2001-07-09
; SEQ ID NO 67
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-09-931-836-67

```

Query Match 100.0%; Score 2399; DB 11; Length 468;

Best Local Similarity 100.0%; Pred. No. 8.7e-150;

Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 MGF LGTGWILVLP IQAFPKPGSQKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
   |||||
DB 1 MGF LGTGWILVLP IQAFPKPGSQKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
   |||||
QY 61 ENKPGQSNYSFVDNLLNKAI TEKEKIERQSRSSPLDNKNLVNVEDVSTKRNKLIDDY 120
   |||||
DB 61 ENKPGQSNYSFVDNLLNKAI TEKEKIERQSRSSPLDNKNLVNVEDVSTKRNKLIDDY 120
   |||||
QY 121 DSTKSGLDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
   |||||
DB 121 DSTKSGLDHKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
   |||||
QY 181 ITESQAHTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIQDGL 240
   |||||
DB 181 ITESQAHTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIQDGL 240
   |||||
QY 241 ARGENDETVSNTLTITNGLERTTYSGEDNFEELQYPNPFVALLKSIDSEKEAKEKETLI 300
   |||||
DB 241 ARGENDETVSNTLTITNGLERTTYSGEDNFEELQYPNPFVALLKSIDSEKEAKEKETLI 300
   |||||
QY 301 TIMKTLDIFVQMVKYGTISPESGVSYLENLDEMIALQTKNLEKNATDNISKLFPAPSE 360
   |||||
DB 301 TIMKTLDIFVQMVKYGTISPESGVSYLENLDEMIALQTKNLEKNATDNISKLFPAPSE 360
   |||||
QY 361 KSHEETDSTKEEAAKMEKEYGSLKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLWK 420
   |||||
DB 361 KSHEETDSTKEEAAKMEKEYGSLKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLWK 420
   |||||
QY 421 KHDKKGKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
   |||||
DB 421 KHDKKGKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
   |||||

```

RESULT 2

US-10-035-977-67

; Sequence 67, Application US/10035977

; Publication No. US20030134327A1

; GENERAL INFORMATION:

; APPLICANT: Desnoyers, Luc

; APPLICANT: Eaton, Dan L.

; APPLICANT: Goddard, Audrey

; APPLICANT: Godowski, Paul J.

; APPLICANT: Gurney, Austin L.

; APPLICANT: Pan, James

; APPLICANT: Stewart, Timothy A.

; APPLICANT: Watanabe, Colin K.

; APPLICANT: Wood, William I.

QY 361 KSHEETDSTKEAAKMEYGLKSTKDDNSNPGKTDPEKGTAEYLAIRKNIWLK 420
 DB 361 KSHEETDSTKEAAKMEYGLKSTKDDNSNPGKTDPEKGTAEYLAIRKNIWLK 420
 QY 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468
 DB 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468

RESULT 3

US-10-063-735-150
 ; Sequence 150, Application US/10063735
 ; Publication No. US2003013882A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Grimaldi, Christopher J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3230R1C1
 ; CURRENT APPLICATION NUMBER: US/10/063,735
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 170
 ; SEQ ID NO 150
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-063-735-150

Query Match 100.0%; Score 2399; DB 12; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8.7e-150;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFLGTGTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
 DB 1 MGFLGTGTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60

QY 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKQIRSPDLNKNLVNEDVDSTNKRKLIDDY 120
 DB 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKQIRSPDLNKNLVNEDVDSTNKRKLIDDY 120

QY 121 DSTKSGLDHFKFQDDPDGLHOLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
 DB 121 DSTKSGLDHFKFQDDPDGLHOLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180

QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTMAAIQDGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTMAAIQDGL 240

QY 241 AKGENDETVSNTLTNLGLERRTKTYSNDFEELQYFPNFPYALLKSIDSEKAKEKETLI 300
 DB 241 AKGENDETVSNTLTNLGLERRTKTYSNDFEELQYFPNFPYALLKSIDSEKAKEKETLI 300

QY 301 TIMKTLIDFVKMVKYGTISPEGVSYLENLDEMIALQTNKLEKNATNISKLPAPSE 360
 DB 301 TIMKTLIDFVKMVKYGTISPEGVSYLENLDEMIALQTNKLEKNATNISKLPAPSE 360

QY 361 KSHEETDSTKEAAKMEYGLKSTKDDNSNPGKTDPEKGTAEYLAIRKNIWLK 420
 DB 361 KSHEETDSTKEAAKMEYGLKSTKDDNSNPGKTDPEKGTAEYLAIRKNIWLK 420

QY 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468
 DB 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468

RESULT 4

US-10-137-870-90
 ; Sequence 90, Application US/10137870
 ; Publication No. US2003013883A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tuma, Daniel
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE OF INVENTION: ACIDS ENCODING THE SAME
 ; FILE REFERENCE: P3330R1C155
 ; CURRENT APPLICATION NUMBER: US/10/137,870
 ; CURRENT FILING DATE: 2002-05-03
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 90
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-137-870-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8.7e-150;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFLGTGTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
 DB 1 MGFLGTGTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60

QY 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKQIRSPDLNKNLVNEDVDSTNKRKLIDDY 120
 DB 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKQIRSPDLNKNLVNEDVDSTNKRKLIDDY 120

QY 121 DSTKSGLDHFKFQDDPDGLHOLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
 DB 121 DSTKSGLDHFKFQDDPDGLHOLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180

QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTMAAIQDGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTMAAIQDGL 240

QY 241 AKGENDETVSNTLTNLGLERRTKTYSNDFEELQYFPNFPYALLKSIDSEKAKEKETLI 300
 DB 241 AKGENDETVSNTLTNLGLERRTKTYSNDFEELQYFPNFPYALLKSIDSEKAKEKETLI 300

QY 301 TIMKTLIDFVKMVKYGTISPEGVSYLENLDEMIALQTNKLEKNATNISKLPAPSE 360
 DB 301 TIMKTLIDFVKMVKYGTISPEGVSYLENLDEMIALQTNKLEKNATNISKLPAPSE 360

QY 361 KSHEETDSTKEAAKMEYGLKSTKDDNSNPGKTDPEKGTAEYLAIRKNIWLK 420
 DB 361 KSHEETDSTKEAAKMEYGLKSTKDDNSNPGKTDPEKGTAEYLAIRKNIWLK 420

QY 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468
 DB 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468

RESULT 5

US-10-140-018-90
; Sequence 90, Application US/10140018
; Publication No. US2003013885A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C158
; CURRENT APPLICATION NUMBER: US/10/140,018
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-018-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150; Indels 0; Gaps 0;
Matches 468; Conservative 0; Mismatches 0;

Qy	1	MGFLGTGTLVLVLP	IQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP	60
Db	1	MGFLGTGTLVLVLP	IQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP	60
Qy	61	ENKPGQSNYSFV	DNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY	120
Db	61	ENKPGQSNYSFV	DNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY	120
Qy	121	DSTKSGLDHKKF	QDDPDGLHQLDGTPLTAEDI VH K I A A R I Y E E N D R A V F D K I V S K L L N I G L	180
Db	121	DSTKSGLDHKKF	QDDPDGLHQLDGTPLTAEDI VH K I A A R I Y E E N D R A V F D K I V S K L L N I G L	180
Qy	181	ITESQAHTLEDE	VAEVLQKLISK EANNYEDPNKPTSWTENQAGK I P E K V T P M A A I Q D G L	240
Db	181	ITESQAHTLEDE	VAEVLQKLISK EANNYEDPNKPTSWTENQAGK I P E K V T P M A A I Q D G L	240
Qy	241	AKGENDET V S N T L T	NG L E R R T K T Y S E D N F E E L Q Y P N F Y A L L K S I D S E K E A K E K E T L I	300
Db	241	AKGENDET V S N T L T	NG L E R R T K T Y S E D N F E E L Q Y P N F Y A L L K S I D S E K E A K E K E T L I	300
Qy	301	TIMKTLD F V K M V K Y G T I	S P E E G V S Y L E N D E M I A L Q T K N K L E K N A T D N I S K L F P A P S E	360
Db	301	TIMKTLD F V K M V K Y G T I	S P E E G V S Y L E N D E M I A L Q T K N K L E K N A T D N I S K L F P A P S E	360
Qy	361	KSHEETDSTKEE	AAKMEKEYGSLK D S T K D N S N P G K T D E P K G T E A Y L E A I R K N I E W L K	420
Db	361	KSHEETDSTKEE	AAKMEKEYGSLK D S T K D N S N P G K T D E P K G T E A Y L E A I R K N I E W L K	420
Qy	421	KHDKKGNKEDY	L S K M R D F I N K Q A D A Y V E K G I L D K E A E A I K R I Y S S L	468
Db	421	KHDKKGNKEDY	L S K M R D F I N K Q A D A Y V E K G I L D K E A E A I K R I Y S S L	468

RESULT 6

US-10-140-274-90

US-10-140-021-90

; Sequence 90, Application US/10140021
; Publication No. US2003013886A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C167
; CURRENT APPLICATION NUMBER: US/10/140,021
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-021-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150; Indels 0; Gaps 0;
Matches 468; Conservative 0; Mismatches 0;

Qy	1	MGFLGTGTLVLVLP	IQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP	60
Db	1	MGFLGTGTLVLVLP	IQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP	60
Qy	61	ENKPGQSNYSFV	DNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY	120
Db	61	ENKPGQSNYSFV	DNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY	120
Qy	121	DSTKSGLDHKKF	QDDPDGLHQLDGTPLTAEDI VH K I A A R I Y E E N D R A V F D K I V S K L L N I G L	180
Db	121	DSTKSGLDHKKF	QDDPDGLHQLDGTPLTAEDI VH K I A A R I Y E E N D R A V F D K I V S K L L N I G L	180
Qy	181	ITESQAHTLEDE	VAEVLQKLISK EANNYEDPNKPTSWTENQAGK I P E K V T P M A A I Q D G L	240
Db	181	ITESQAHTLEDE	VAEVLQKLISK EANNYEDPNKPTSWTENQAGK I P E K V T P M A A I Q D G L	240
Qy	241	AKGENDET V S N T L T	NG L E R R T K T Y S E D N F E E L Q Y P N F Y A L L K S I D S E K E A K E K E T L I	300
Db	241	AKGENDET V S N T L T	NG L E R R T K T Y S E D N F E E L Q Y P N F Y A L L K S I D S E K E A K E K E T L I	300
Qy	301	TIMKTLD F V K M V K Y G T I	S P E E G V S Y L E N D E M I A L Q T K N K L E K N A T D N I S K L F P A P S E	360
Db	301	TIMKTLD F V K M V K Y G T I	S P E E G V S Y L E N D E M I A L Q T K N K L E K N A T D N I S K L F P A P S E	360
Qy	361	KSHEETDSTKEE	AAKMEKEYGSLK D S T K D N S N P G K T D E P K G T E A Y L E A I R K N I E W L K	420
Db	361	KSHEETDSTKEE	AAKMEKEYGSLK D S T K D N S N P G K T D E P K G T E A Y L E A I R K N I E W L K	420
Qy	421	KHDKKGNKEDY	L S K M R D F I N K Q A D A Y V E K G I L D K E A E A I K R I Y S S L	468
Db	421	KHDKKGNKEDY	L S K M R D F I N K Q A D A Y V E K G I L D K E A E A I K R I Y S S L	468

GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3330R1C174
; CURRENT APPLICATION NUMBER: US/10/140,807
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-807-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGFLGTGTVLVLPIQAFKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTVLVLPIQAFKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVDNLNLLKAITKEKIEKQSRSSPLDNKLNVEDVDSTNKRKLDDY 120
Db 61 ENKPGQSNYSFVDNLNLLKAITKEKIEKQSRSSPLDNKLNVEDVDSTNKRKLDDY 120
Qy 121 DSTKSGLDHKFDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFDKIVSKLLNLGL 180
Db 121 DSTKSGLDHKFDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFDKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIOQGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIOQGL 240
Qy 241 AKGENDETSTNTLTITNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETSTNTLTITNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLIDFVKMWKYGTTISP EEGSVYLENDLMIALQTNKLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVKMWKYGTTISP EEGSVYLENDLMIALQTNKLEKNATDNISKLPAPSE 360
Qy 361 KSHETDSTKEEAAKMEKEYGSLKSDTSDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
Db 361 KSHETDSTKEEAAKMEKEYGSLKSDTSDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKNGKEDYDLKSMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
Db 421 KHKKGKNGKEDYDLKSMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468

RESULT 10
US-10-140-922-90
; Sequence 90, Application US/10140922
; Publication No. US20030138889A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

APPLICANT: Baker, Kevin P.
APPLICANT: Beresini, Maureen
APPLICANT: DeForge, Laura
APPLICANT: Desnoyers, Luc
APPLICANT: Filvaroff, Ellen
APPLICANT: Gao, Wei-Qiang
APPLICANT: Gerritsen, Mary E.
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Gurney, Austin L.
APPLICANT: Sherwood, Steven
APPLICANT: Smith, Victoria
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Watanabe, Colin K
APPLICANT: Wood, William
APPLICANT: Zhang, Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3330R1C179
CURRENT APPLICATION NUMBER: US/10/140,922
CURRENT FILING DATE: 2002-05-07
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 90
LENGTH: 468
TYPE: PRT
ORGANISM: Homo Sapien
US-10-140-922-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGFLGTGTVLVLPIQAFKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTVLVLPIQAFKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVDNLNLLKAITKEKIEKQSRSSPLDNKLNVEDVDSTNKRKLDDY 120
Db 61 ENKPGQSNYSFVDNLNLLKAITKEKIEKQSRSSPLDNKLNVEDVDSTNKRKLDDY 120
Qy 121 DSTKSGLDHKFDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFDKIVSKLLNLGL 180
Db 121 DSTKSGLDHKFDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFDKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIOQGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIOQGL 240
Qy 241 AKGENDETSTNTLTITNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETSTNTLTITNGLERRTKTYSEDNFEELQYFPNFYALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLIDFVKMWKYGTTISP EEGSVYLENDLMIALQTNKLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVKMWKYGTTISP EEGSVYLENDLMIALQTNKLEKNATDNISKLPAPSE 360
Qy 361 KSHETDSTKEEAAKMEKEYGSLKSDTSDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
Db 361 KSHETDSTKEEAAKMEKEYGSLKSDTSDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKNGKEDYDLKSMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
Db 421 KHKKGKNGKEDYDLKSMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468

RESULT 11
US-10-140-924-90
; Sequence 90, Application US/10140924
; Publication No. US20030134355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

APPLICANT: Gao,Wei-Qiang
APPLICANT: Gerritsen,Mary E.
APPLICANT: Goddard,Audrey
APPLICANT: Godowski,Paul J.
APPLICANT: Gurney,Austin L.
APPLICANT: Sherwood,Steven
APPLICANT: Smith,Victoria
APPLICANT: Stewart,Timothy A.
APPLICANT: Tumas,Daniel
APPLICANT: Watanabe,Colin K
APPLICANT: Wood,William
APPLICANT: Zhang,Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3330RIC209
CURRENT APPLICATION NUMBER: US/10/141,704
CURRENT FILING DATE: 2002-05-08
Prior Application removed - See Palm or File Wrapper
NUMBER OF SEQ ID NOS: 550
SEQ ID NO 90
LENGTH: 468
TYPE: PRT
ORGANISM: Homo Sapien
US-10-141-704-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGWILVLPVLPQAPKPGSQDSKSHNELSAERPLNEQIAEAEDKIKKTYPP 60
DB 1 MGFLGTGWILVLPVLPQAPKPGSQDSKSHNELSAERPLNEQIAEAEDKIKKTYPP 60
QY 61 ENKPGQSNYSFVDNLNLLKATEKEKIERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
DB 61 ENKPGQSNYSFVDNLNLLKATEKEKIERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
QY 121 DSTKSGLDHFKPDPPDGLHQLDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGL 180
DB 121 DSTKSGLDHFKPDPPDGLHQLDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGL 180
QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKPEKVTNMAAIODGL 240
DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKPEKVTNMAAIODGL 240
QY 241 AKGENDETVENTLTNGLERRTKTYSEDNFEBLQYFPNPFYALLKSIDSEKAKEKETLI 300
DB 241 AKGENDETVENTLTNGLERRTKTYSEDNFEBLQYFPNPFYALLKSIDSEKAKEKETLI 300
QY 301 TIMKTLIDFVKMMVKYGTISPEGVSVLENLDEMIALQTKNLEKNATDNISKLPFPAPSE 360
DB 301 TIMKTLIDFVKMMVKYGTISPEGVSVLENLDEMIALQTKNLEKNATDNISKLPFPAPSE 360
QY 361 KSHHEETDSTKEAAKMEKEYGSLKSDKDNNSPGGKTDEPKGKTEAYLEAIRKNIEWLK 420
DB 361 KSHHEETDSTKEAAKMEKEYGSLKSDKDNNSPGGKTDEPKGKTEAYLEAIRKNIEWLK 420
QY 421 KHDKGNKEDYDLCKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
DB 421 KHDKGNKEDYDLCKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468

Search completed: January 2, 2004, 12:37:45
Job time : 40 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 12:33:21 ; Search time 20 Seconds
(without alignments)
2250.347 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGTWLVLPQAF.....EKGILDKEEBAIKRIYSSL 468
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues
Total number of hits satisfying chosen parameters: 283308
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries
Database : PIR_76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1957	81.6	533	2 A37180	chromogranin/secre
2	196	8.2	2139	2 T18296	myosin heavy chain
3	191	8.0	2464	1 QMSPL	microtubule-associ
4	186.5	7.8	2116	2 A26655	myosin heavy chain
5	184	7.7	665	2 B71609	hypothetical prote
6	183.5	7.6	1790	2 S67593	transport protein
7	180.5	7.5	2364	2 A56577	microtubule-associ
8	179.5	7.5	1526	2 A45605	mature-parasite-in
9	179	7.5	853	2 T51505	hypothetical prote
10	179	7.5	1005	2 A64485	hypothetical prote
11	179	7.5	1558	2 B71603	RESA-H3 antigen PF
12	179	7.5	2269	2 T28677	rhostry protein -
13	178.5	7.4	1392	2 A43336	microtubule-vesicl
14	178.5	7.4	1427	2 S22695	restin - human
15	177	7.4	1871	2 D36796	probable heat shoc
16	172.5	7.2	1134	2 A60234	IGA Fc receptor pr
17	172.5	7.2	1164	1 FCSOAG	IGA Fc receptor pr
18	172.5	7.2	1979	2 C71632	hypothetical prote
19	171.5	7.1	3488	2 T34418	hypothetical prote
20	170	7.1	1804	2 T34518	neatin - golden ha
21	169	7.0	1272	2 C30593	hypothetical prote
22	168.5	7.0	746	2 T47237	myosin II heavy ch
23	168.5	7.0	926	2 T24923	hypothetical prote
24	168	7.0	2401	2 T28676	rhostry protein -
25	167	7.0	976	2 T01553	hypothetical prote
26	166.5	6.9	1875	2 S38173	myosin-like protei
27	166.5	6.9	5327	2 T13564	microtubule-associ
28	166	6.9	991	2 H86168	hypothetical prote
29	166	6.9	1192	2 A71623	probable secreted

RESULT 1
A37180
Chromogranin/secretogranin-like vesicle protein precursor - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 30-Sep-1993
C;Accession: A37180
R;Otter, H.P.; Battenberg, E.F.; Tsou, A.P.; Bloom, F.E.; Sutcliffe, J.G.
J. Neurosci. 10, 3135-3147, 1990
A;Title: B1075: a brain- and pituitary-specific mRNA that encodes a novel chromogranin/
A;Reference number: A37180; MUID:90376160; PMID:2204688
A;Accession: A37180
A;Status: preliminary; not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-533 <OTT>

Query Match 81.6%; Score 1957; DB 2; Length 533;
Best Local Similarity 86.7%; Pred. No. 3.5e-87;
Matches 383; Conservative 21; Mismatches 38; Indels 0; Gaps 0;

Qy	16	PIQAFKPGGSDKSLHNRELSAERPLNEQIAEAEDKIKKTPPENKPGQSNYSFVDNL	75
Db	63	PNSFPKPGGSDKSLHNRELSAERPLNEQIAEAADKIKKTPSEKSPERNFSSVDNL	122
Qy	76	NLLKAITEKEKIEKQSISSPLDNKLNVEDVDSTKNRKLIDDYDSTKGLDHFQDDP	135
Db	123	NLLKAITEKETVEKAKQSISSPFDNLDVDDADSTKNRKLTDYDSTKGLDKVQDDP	182
Qy	136	DGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLNLGLITESQAHTLEDEVAE	195
Db	183	DGLHQLDGTPLTAEDIVHKIATRIYEENDRGVFDKIVSKLNLGLITESQAHTLEDEVAE	242
Qy	196	VLOKLISKEANNYEEDPNKPTSTENQAGKIPKVTFMAAIOGLAKGENDETSTNLTLL	255
Db	243	ALQKLISKEANNYEAEPEKPTSTENQDGKIPKVTFAATQDGTNRENDTSTNLTLL	302
Qy	256	TNGLERTKTYSDNFEELQYFNFVALLKSIIDSEKAEKETLITIMKTLIDFVKMVK	315
Db	303	SNGLERTNPHRDDFEELQYFNFVALLTSIDSEKAEKETLITIMKTLIDFVKMVK	362
Qy	316	YGTISPEEGVSYLENLDEMIALQTKNLEKNAIDNLSKLPAPSEKSHETDSTKEBAK	375
Db	363	YGTISPEEGVSYLENLDETALQTKNLEKNTDTSKSLFPAPSEKSHETDSTKEBAK	422
Qy	376	MEKEYSLKSDTKDDNSNPGKTDDEPKGKTEAYLEAIRKNIEMLKDKKGNKEDYLSK	435
Db	423	MEKEYSLKSDTKDDNSNLGKTKDEAKGKTEAYLEAIRKNIEMLKDKKGNKEDYLSK	482
Qy	436	MRDFINKQADAYVEKGILDKEE	457
Db	483	MRDFINQADAYVEKGILLIRKK	504

centromere protein
hypothetical prote
hypothetical prote
ORF MSV156 hypothe
probable myosin he
liver stage antige
reticulocyte-bindin
hypothetical prote
interaptin - slime
myosin heavy chain
myosin heavy chain
surface-located me
myosin heavy chain
conserved hypothet
serine/threonine-s
hypothetical prote

T18296
myosin heavy chain - Entamoeba histolytica
C:Species: Entamoeba histolytica
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 08-Sep-2000
C:Accession: T18296
R:Guillen, N.
submitted to the EMBL Data Library, February 1997
A:Reference number: Z18865
A:Accession: T18296
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2139 <GUI>
A:Cross-references: EMBL:L03534; NID:g1850912; PID:g1850913; PIDN:AAB48065.1
C:Genetics:
A:Gene: mhca
C:Superfamily: myosin heavy chain; myosin motor domain homology
F:91-780/Domain: myosin motor domain homology <MMO>

Query Match 8.2%; Score 196; DB 2; Length 2139;
Best Local Similarity 21.0%; Pred. No. 0.1;
Matches 107; Conservative 108; Mismatches 191; Indels 104; Gaps 23;

Qy 20 FPKPGSGDKSLHNRSLSAERPLNEQTAAEAEDKIKTYPEN-----KPGQSNTSVFV 72
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 859 FEPEGKKDKIED-----LKKKLAE-----EIKREAAENALASATANKGLEAKIQ 906

Qy 73 D-----NLNLKAITEKEKIEKERQSIRSSPLDNKLNAVVDV- 109
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 907 DLEKDISESKLSAAELDKQLNLKTLENLEEDKEELKETIDNLKGDLKSKLGEDLEV 966

Qy 110 -----STKNRKLLDDYSTKSGLDHKPDPPDGLHQLDGTPLTFAEDIVHKAARIY 160
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 967 EITELNSQINTLN--ATWDKKDKTTIAEQMESIDEKEDEITTKLGDIKLEBEKODL---- 1020

Qy 161 EEENDR-----AVFDKIVSKLNLGLITISOAHTLDEVAEVLOKLI SKEANNYEEDPNKPT 216
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 1021 -EQRADRVSATKDIIAKLNKITI-----ECEDAKDEIAKUEQELEDSE--NNKNDLTNEL 1073

Qy 217 SWTENQAGKIPEKVTPMAAI-----QDGLAKG-ENDETVSNTLTITNG-LERRTKTYS 267
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 1074 QOTQLKLGETEKSLAAQAATKASKASDERDTLSQNLENEKLTNNLTTKADLEKKISGLK 1133

Qy 268 EDNFEELOYFPNFALLKSID-----SEKAKEKETIITIMKTILIDPV-KMVKYGTISP 321
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 1134 QD-YEDLEDDKN-----KIEGLDRNAQRKIKELDDBITKGADVSYLOKQKEBY----- 1181

Qy 322 EEGVSYLENDEMIALQTNKLNKATNTSILKLPAPSEKSHEETDSTKEAAK-----ME 377
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 1182 ESQIAKMQEKEKALGNDVNKN-EKTIKE--KELEIQLQELDETEVEKESDAEKKKKEIE 1238

Qy 378 KEYGSGLKDYTDNSNPGGKTDSPFKGTEAYLAIRKNIEWLKHDKKGNKEDYDL-SKM 436
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 1239 KEMKALQEE-KENVESSKNSTEKKKKLEDNLDKDTOKKLLDMWTADNEKLKAKAKDLEAQL 1297

Qy 437 RDPFNKQAQAVYKVGILDKEEAIAKRIYS 466
Db :|::||:::||||:||||:||||:||||:||||:||||:||||:||||:||||:|
Db 1298 NEVQDNHKEAVADAELLNKKKAQSDKELNS 1327

RESULT 3
QRMSP1
microtubule-associated protein MAP1B - mouse
N:Alternate names: microtubule-associated protein MAP1(X); microtubule-associated protein
C:Species: Mus musculus (house mouse)
C>Date: 30-Sep-1991 #sequence_revision 30-Sep-1991 #text_change 01-Sep-2000
C:Accession: S07549; S44387; A33645
R:Noble, M.; Lewis, S.A.; Cowan, N.J.
J. Cell Biol. 109, 3367-3376, 1989.
A>Title: The microtubule binding domain of microtubule-associated protein MAP1B contains
A:Reference number: A33645; MUID:90094539; PMID:2480963
A:Accession: S07549
A:Molecule type: mRNA

Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
A>Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
A/Reference number: A26655; MUID:87092266; PMID:3540939
A/Accession: A26655
A/Molecule type: DNA
A/Residues: 1-2116 <WAR>
A/Cross-references: GB:M11938; NID:g167834; PIDN:AAA33227.1; PID:g167835
R:DeLozanne, A.; Lewis, M.; Spudis, J.A.; Levinwand, L.A.
Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
A/Reference number: A24728; MUID:86016788; PMID:3901008
A/Accession: A24728
A/Molecule type: mRNA
A/Residues: 2035-2116
R:Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
FEBS Lett. 227, 71-75, 1988
A>Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium myosin heavy chain
A/Reference number: S00250; MUID:88112226; PMID:2828113
A/Accession: S00250
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1734-1893 <WAG>
A/Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphate binding site
F:1-818/Domain: Globular head <HED>
F:89-747/Domain: myosin motor domain homology <WMOT>
F:179-186/Region: nucleotide-binding motif A (P-loop)
F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 7.8%; Score 186.5; DB 2; Length 2116;
Best Local Similarity 21.6%; Pred. No. 0.29;
Matches 108; Conservative 86; Mismatches 205; Indels 101; Gaps 17;
Db 1023 TKDSKSELRLQKKLEELKQVQALAEATAKLAQEAANKLQGEYTEL-NEKFNSEVTA 1081
Qy 84 KEKIEKERQSTRSS--PLDNKLVNEDVDSTNKRKLIDYDSTKSGLDHKKFDDPGLHLQ 141
Db 1082 RSNVEKSKTLESQVAVNNELDEE---KQNR---DALEKKKALDAMLEMKDQLEST 1134
Qy 142 DGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGLITESQAHTLEDEVAEVLKLI 201
Db 1135 GGEKSLYDL-----VKQSDMEALRNQISELQSTIAKLEKTKSTLEGEVRLQGELE 1188
Qy 202 SKK--ANNVEEDPNKPTSWTNOAGKIPEKVT-----PMAAQDGLAKEN 245
Db 1189 AEQLAKSNVEKQKKVVELDLEKSKQALAEATAKQALDKLKKLEQELSEVQTOLSEANN 1248
Qy 246 D-----ETVSNLTITLNGLERTKTYSEDN----FEELYFPNFPYALLKSIDSE 290
Db 1249 KVNNSDSTNKHLETSFNKLLEAEQAKQALEKRLGLESELKH-----VNEQLEEE 1302
Qy 291 KEAKEK-----ETLITIMKTLDIFKVMVKVGTITSPSEGVSYLE----- 329
Db 1303 KQKESNEKRVKDLKEVSELSKQDTEEVASKAVTEAKNKKSELSDELIRKQYADVSSR 1362
Qy 330 --NIDEMIALOTQN-KLEKNATDNISKLFPAPSEKSHETD-----STKEAAKMEKEY 380
Db 1363 DKSVEQLTKLQAKNEELNTEAEAGQDLRAERSKKAFFLEAVKNLEETAKKVAE 1422
Qy 381 GSKDKSTKDDNSFGGKTDEPKGKTEAYLEAIRKNIEWLKHKHKKGNKEDYDLSKMRDFI 440
Db 1423 KAMKKAETDYRSTKSELDDAKNVSSEQVQIKRLNEE-----LSLSRLV 1467
Qy 441 NKQADAYVEKILDKKEAEA 460
Db 1468 -EADERCNSAIKAKTKTAES 1486

RESULT 5
B71609
hypotheical protein PFB0680w - malaria parasite (Plasmodium falciparum)

C/Species: Plasmodium falciparum
C/Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C/Accession: B71609
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
Perte, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
Science 282, 1126-1132, 1998
A>Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A/Reference number: A71600; MUID:99021743; PMID:9804551
A/Accession: B71609
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-665 <GAR>
A/Cross-references: GB:AE001410; GB:AE001362; NID:g3845245; PIDN:AACT1925.1; PID:g3845245
A/Experimental source: clone 3D7
C/Genetics:
A/Gene: PFB0680w

Query Match 7.7%; Score 184; DB 2; Length 665;
Best Local Similarity 22.0%; Pred. No. 0.089;
Matches 104; Conservative 76; Mismatches 182; Indels 110; Gaps 17;
Qy 25 GSODKSLHNRSLSAERPLNQIAEAEDKIKTYPENKPCQSNYSFVDNLLNLKAITEK 84
Db 160 GKQDISNSNAE--NKQDVKGKVELEKKEEKISDDHKVEENKKS-----DD 205
Qy 85 EKIEKERQSTRSSPLDNKLN---VEDVDSTNKRKLIDYDSTKSGLDHKKFDDPGLHLQ 140
Db 206 HKVEENKKSDDHKVEENKSDHKKIEEVKVEEHEDEEDKKEKSENKKNKDNKDEND 265
Qy 141 LDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGLITESQAHTLEDEVAEVLKLI 200
Db 266 EDNDEISDEDDVDVEEDKNENDIDDDK-----KETDKTHLEEEENIEEKE 314
Qy 201 IS-KEANNYEDPNKPTSW-TENQAGKIPEKVTMAAIQDGLAKGENDETNTLTITNG 258
Db 315 FSDKKGKKNKDKYKESKDKTEKSKDIEK-----EKSQKEKESK----- 356
Qy 259 LERTTKTYSEDNFEELYFPNFPYALLKSIDSEKAEKETLITIMKTLDIFVGMVKYGT 318
Db 357 -KQKEKKGDKKE-----KSKDIEK-KKQDKDIEKSK-DTAKEKEKDKD 402
Qy 319 ISPEGVSYLENLDEMIALQTKKLEKATNTNISKLPAPSEKSH-----E 364
Db 403 IEKESK-----KQKELKKNQKDEKDKDNEKK-----KNDKQIDHDDNDDEMEIE 451
Qy 365 ETSTKEAAKME-----KEYGSLKSTPK-DNNSNPGGKTDEPKGKTEAYLEAI 412
Db 452 ENDEDEDEDMENKKKKKGKNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGN 511
Qy 413 RKNIEWLKRDHKKGNKEDYDLSKMRDF-----INKQADAYVEKGLID 454
Db 512 NENNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGN 563

RESULT 6
S67593
transport protein US01 - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein D2552; protein YDL058w
C/Species: Saccharomyces cerevisiae
C/Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
A/Accession: S67593; A38455; S30782
R:Blöcker, H.; Brandt, P.
submitted to the Protein Sequence Database, July 1996
A/Reference number: S67587
A/Accession: S67593
A/Molecule type: DNA
A/Residues: 1-1790 <BLO>
A/Cross-references: EMBL:274106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058w
A/Experimental source: strain S288C
R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamaeaki, M.
J. Cell Biol. 113, 245-260, 1991
A>Title: A cytoskeleton-related gene, US01, is required for intracellular protein transp
A/Reference number: A38455; MUID:91185402; PMID:2010462

A:Accession: A38455
 A:Molecule type: DNA
 A:Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>
 A:Cross-references: GB:X54378; NID:64777; PIDN:CAA3253.1; PID:94778
 A:Note: the authors translated the codon ACT for residue 768 as Ile
 R:Hostetter, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.
 submitted to the EMBL Data Library, February 1993
 A:Description: An integrin analogue in *Saccharomyces cerevisiae*.
 A:Reference number: S30782
 A:Accession: S30782
 A:Molecule type: DNA
 A:Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'
 A:Cross-references: EMBL:L03188
 C:Genetics:
 A:Gene: SGD:USO1; INT1
 A:Cross-references: SGD:S0002216; MIPS:YDL058W
 A:Map position: 4L
 C:Keywords: coiled coil; transmembrane protein
 F:326-342/Domain: transmembrane #status predicted <TM1>
 F:394-410/Domain: transmembrane #status predicted <TM2>
 F:617-633/Domain: transmembrane #status predicted <TM3>

Query Match 7.6%; Score 183.5; DB 2; Length 1790;
 Best Local Similarity 20.3%; Pred. No. 0.33;
 Matches 120; Conservative 91; Mismatches 200; Indels 179; Gaps 23;

QY 29 KSLHN--RELAEPLNQLAEAEEDKIKTYPPENKPGQSNYSFVDNLLKAI---T 82
 Db 965 KSLANNYKDMQAE--NESLIKAVEE-----SKVSSIQSLNQLKIDMSQ 1008

QY 83 EKEKIEKERSQIRSPNDKLNVEDVSTQVKLI-----DYDSTKSGLDHKFPD--- 133
 Db 1009 EKENQIERGSIKRIEQLKTKTIDLEQTKBEIISKSDSKDEYESQISLLKEKLETATT 1068

QY 134 -DPDLGLDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGLITESQAH----- 187
 Db 1069 ANDENVNKKISLTITRELEAEALAA--YKNLKNELETETSEKALKEVNEBEHLKKEK 1126

QY 188 -TLDEVAEVLQKLIKEAN-----NYEED-PNKPTSWTENQAGKIPEK 229
 Db 1127 IQLEKEATETKQLNSLRANLESEKHEHDLAQLKKYEQIANKGRQYNE-EISQINDE 1185

QY 230 VTPMAAIDGLAKGENDETVSNLTITNGLERRT-----KTYSDN-- 270
 Db 1186 ITSTQOENESIKK-KNDELGEVAMKSTSEOSNLKKSIFDALNLQIKELKKKNETNEA 1244

QY 271 -----FEEQLQFPNF-----YALLKSTIDSEKAK 294
 Db 1245 SLLESIKSVESETVKIKELQDECNFKEVSELEDKIKASEDKNKLQK-ESEKIKE 1303

QY 295 EKETLITIMKTLDIFVKQMWK----- 315
 Db 1304 ELDAKTTTELQIKELKITNLKAKEKSESELSRLKKTSSBERKNAERQLEKLNKNEIQIKNQ 1363

QY 316 -----YGTISPE--EGSVYLENDEMIALQTKNKLKKNATDNISKLPAPSEK 361
 Db 1364 AFEKERKLLNEGSSITTOEYSEKINILE--DEIRLQENENELKAKEDINTRSELEKVSLS 1421

QY 362 SHBETOSTKEAAKMEKYEKSLKDS-TKDDNSNPGGKTDEP-----GKTEAYLEAIRK 414
 Db 1422 NDELLSEKQNTIKSLQDEILSYKDKITRNDKLLSTERNKRDLSLEKQLRAAQESKAK 1481

QY 415 NIEWLAKHDKKNGEDYDLSKMRDFINKQADAYVEKGLDKEAE 463
 Db 1482 VEELKLEESSEKAELEKSEKEM--KKLESTIESNETELKSSMETIRK 1530

RESULT 7
 A56577
 microtubule-associated protein MAP 1B - rat (fragment)
 C:Species: Rattus norvegicus (Norway rat)
 C>Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Feb-1997
 C:Accession: A56577

R:Zauner, W.; Kratz, J.; Staunton, J.; Feick, P.; Wiche, G.
 Eur. J. Cell Biol. 57, 66-74, 1992

A:Title: Identification of two distinct microtubule binding domains on recombinant rat M
 A:Reference number: A56577; MUID:92347374; PMID:1639092

A:Accession: A56577

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-2364 <ZAU>

A:Cross-references: GB:X60550

A:Experimental source: brain

A:Note: nucleotide sequence not given; conceptual translation not complete

C:Superfamily: microtubule-associated protein MAP1B

Query Match 7.5%; Score 180.5; DB 2; Length 2364;
 Best Local Similarity 22.7%; Pred. No. 0.65;
 Matches 120; Conservative 77; Mismatches 200; Indels 131; Gaps 22;

QY 22 KPGGSQDKSLHNRLELSAERPLNEQIAEAEEDKIK-KTYPPENKPGQSNYSFVDNLLKLA 80
 Db 454 KFLSSKSVKESKEAPEATKASQVEKTPKVESKEKVIKKDKPGK-----VESKPS 505

QY 81 ITEKIEKERSQIRSPNDKLNVEDVSTKNRKLIDDYDSTKSGLDHKFODDP--DGL 138
 Db 506 VTEKEVPSKEEQSPVKAEEAATESPKVTCKDKVVKKEIKTKP--EKKKEKPKKEVA 563

QY 139 HOLDGTPLTAEDIVHKIARIYEENDRAVDFKIVSKLNLGLITESQAH--LEDEVAEVL 197
 Db 564 KKEDKTPKDKDEKPKK-----EPAKKEIKKEIKKEKELKKEVKKETPLKDAKKEV- 615

QY 198 QKLIKEKANNVEEDPNK-----PTSWTENQAG-----KIPKVTTPMAA- 235
 Db 616 KDKENKVKKEKEPKKEIKKIDIKKSTPLSDTKPKPAALKPKVAKKEETKKEPIAAG 675

QY 236 -IQD-GLAK-----GENDETVSNTL-----TLTNGLERRTK 264
 Db 676 KLDKRGKVKVKKEGKTTTAAATAVGTAAVAAAAGVAASGPAKELAEBSLWSSPEDITK 735

QY 265 TVSEDNFEEQLQFPNFYALLKSIDSEKAE-----KETLITIMKTLDIFVKQMW 314
 Db 736 DFEEKLAKEIDVAKDIPQLSLIDEKELKETEPEGEAYVIQKETEVS----- 782

QY 315 KYGTISPEGVSVYLENDEMIALQTKNKL-----EKNATDNISKL-----FPAPSEKSHSE 365
 Db 783 KGSAESPDGEGITTEGECE--CEQTPPELEPVEKQGVDDIEKFEDEGAGFESEAGDYE 840

QY 366 TDSITKEEAAKMEKYEKSLKDS-TKDDNSNPGGKTDEPKGKTEA--YLBAIRNI----- 416
 Db 841 EKAETEAEAEPEEDG---EDNVSGSASKHSPTDEEETAKAEADVHIKEKRESVASGDRA 897

QY 417 -----EWLKKHDKKNGEDYDLSKMRDFINKQADAYVEKGLDKEAE 459
 Db 898 EEDWDEALEKGEAQSEEGEEEEE-----DKAEDAREEDHEPDKTEAE 940

RESULT 8

A56005

mature-parasite-infected erythrocyte surface antigen MESA - malaria parasite (Plasmodium

C:Species: Plasmodium falciparum

C>Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-2000

C:Accession: A45605; A54517

R:Coppel, R.L.

Mol. Biochem. Parasitol. 50, 335-347, 1992

A:Title: Repeat structures in a Plasmodium falciparum protein (MESA) that binds human ery

A:Reference number: A45605; MUID:92158014; PMID:1741020

A:Accession: A45605

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-1526 <COP>

A:Experimental source: Papua New Guinean isolate FC27

A:Note: sequence extracted from NCBI backbone (NCBI:83648, NCBI:83656)

R:Coppel, R.L.; Culvenor, J.G.; Bianco, A.E.; Crewe, P.E.; Stahl, H.D.; Brown, G.V.; J

Mol. Biochem. Parasitol. 20, 265-277, 1986

A:Title: Variable antigen associated with the surface of erythrocytes infected with matu

A:Reference number: A54517; MUID:87014571; PMID:3531849

A:Accession: A54517

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 222-443, 'K', 445 <C02>

A:Cross-references: GB:M15319; NID:g160060; PID:g552170

C:Superfamily: ring-infected erythrocyte surface antigen; dnaJ amino-terminal homology
C:Keywords: surface antigen; tandem repeat

Query Match 7.5%; Score 179.5; DB 2; Length 1526;
Best Local Similarity 18.7%; Pred. No. 0.42;
Matches 107; Conservative 101; Mismatches 182; Indels 181; Gaps 24;

Qy 25 GSQKSLNRLS--AERPLNEQTAABEDKIKTYPPENKPGQSN----- 68

Db 400 GASSENSEDPKLTQEENGTGTSSEETKDDK-----PEENEKADNKKSKKKKSPFQM 454

Qy 69 --YSPVDNLN-----LKAITEKIEKERSIRSSPLDNKLVNEDVDSTNKRKLIDDYD 121

Db 455 LGCNPLCNKNIETDDEETLVVKODAKKKHFLREA-----NTEKNDNEKKKKLGGD 508

Qy 122 -----STKSGLDHKFQDDPDGL-----HOLDGTPL 146

Db 509 KEDVKKNDEQKVLGSGDKEDVKEKNDQKDLVGBGDKEDVKEKNDGKKKGVISEK 568

Qy 147 TAEDIVHKIAARI-----YEENDRAVFKIVSKLNLGLITESQAHITLEDV-- 193

Db 569 TQKEIKERVKRRVKKCKKKVKKGKENDTEGNDKVGPEIIEBVKEEIKKQVEDGIKE 628

Qy 194 -----AEVLQKLSKEANNYEDPNKPTSWTENQAGKIPKVT-----MAAI 236

Db 629 NDTGNDKVGPEIIEBVKEEIKKQVEDGKENDTEGNDKVGPEIIEBVKEEIKKQV 688

Qy 237 QDGLAG--ENDETNSNTLTITNGLERRTK-----TYSEDNF-----EELQY 276

Db 689 EEGIKENDTEGNDKVGPEIIEBVKEEIKKQVEGVKENDTESKDKVIGOEIITEEVK- 747

Qy 277 PNFYALKSIDSKEAKEKETLITI-----MTLIDFVQMVKYGTISPGEV--SYL 328

Db 748 -----KETQBEKNGKENILEIKDIVIGQEVIIIEVVKVIK--KVEKGIKENHT 796

Qy 329 ENLD-----EMALQTKNKKLNATDNISKLPAPSEK-----SHEETDSTKEA 373

Db 797 ESKDKVIGOEIIEBVKEEIEKQVEGVKENDTESKDKVIGEVKGVNBEVSGPENKDKV 856

Qy 374 AKMEKEYGSLKSDTKDDNSNPGKTDPEKGTAEYLAIRKNIEMWLKHKDKKNGKEDYDL 433

Db 857 TKQEK--VKEVK-----EVKKK-----VKRV--KRNKNKRNKONVI 891

Qy 434 SK--MRDPINKQADAYVEKGILDKKEEAPAK 462

Db 892 GKEIMKEDVNEKTANKDKETEQSEKEEVK 922

RESULT 9

T51505

hypothetical protein F5E19_70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 18-Aug-2000 #sequence_revision 18-Aug-2000 #text_change 18-Aug-2000

C:Accession: T51505

R:Sato, S.; Nakamura, Y.; Kaneko, T.; Kato, T.; Asamizu, E.; Kotani, H.; Tabata, S.; Mew

submitted to the Protein Sequence Database, August 2000

A:Reference number: 225394

A:Accession: T51505

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-853 <SAT>

A:Cross-references: EMBL:AL391147

A:Experimental source: cultivar Columbia; BAC clone F5E19

C:Genetics:

A:Map position: 5

A:Introns: 6/2; 79/3

A:Note: F5E19_70

Query Match 7.5%; Score 179; DB 2; Length 853;

Best Local Similarity 20.9%; Pred. No. 0.21;

Matches 102; Conservative 81; Mismatches 206; Indels 100; Gaps 19;

Qy 16 PIQAFPPGSGSQSLNRLSAR-----PLNEQIAEAEEDKIKTYPPEN----- 62

Db 45 PSTTTPHSRLSLDRSSNSKSSVERRSPKLTTPPEKSOARVAAVKGTESPTTTLRSLQIK 104

Qy 63 ---KPGQSNYSFVDNLNLLKAIKTEKIEKERSIRSSPLDNKLVNEDVDSTNKRKLIDD 119

Db 105 EDLKKANERISSLEK-DKAKALDELKQAKAEQV-TLKLDDALKAQ--KHVEENSIEIK 160

Qy 120 YDSTKSGLDHKFQDDPDGL-----HOLDGTPLTAEDIVHKIAARIYEENDRAVFD 169

Db 161 FQAVEAGIE-AVQNNBEELKKELETVKQKHASDAALVA---VRQELEKINEEL-AAAFD 215

Qy 170 KIVSKLNLGLITESQAHITLEDVAEVLQKLSKEANNYEDPNKPTSWTENQAGKIPK 229

Db 216 -AKSKAL-----SQA-----EDASKTAIEIHAERKVDILSE 244

Qy 230 VTPMAAIQDGLAKGENDETVNSLTITNGLERRTKTYSEDNFEELOYPNFPYALKKIDS 289

Db 245 LTRLKALLD-----STREKTAISDNEMVAKLEDEIVVLKRDES 283

Qy 290 ----EKEAKEKETLITIMKTLIDFVQMVKYGTISPGEVSYLENDEMIALQTKNLEK 345

Db 284 ARGFEAEVKEKEMIVKLVNDELAAKMAESNAHSLSNMQSKAKELSEQ--LEEANKLER 341

Qy 346 NAT----DNISKLPAPSEKSHETDST--KEAAKMEKEYGSLKSDTKDDNSNPGKTD 399

Db 342 SASVLSVNMVQLEGNDKLHDTETITDLKERIVTLTTVAQKQEDLEVSEQR--GSVE 400

Qy 400 EPKGKTEAYLEAIRKNIEMWLKHKDKKNGKEDYDLKMRDFINKQADAYVEKGILDKKEAE 459

Db 401 EEVSQKNEKVEKLSKSELETVEKKNRALKKEQATSRVQLSEBKSLDLESSKEEEE 460

Qy 460 AIKRIYSSL 468

Db 461 KSKKAMESL 469

RESULT 10

A64465

hypothetical protein MJ1322 - Methanococcus jannaschii

C:Species: Methanococcus jannaschii

C:Date: 13-Sep-1996 #sequence_revision 13-Sep-1996 #text_change 21-Jul-2000

C:Accession: A64465

R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blake, C.

; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, K.G.; Merrick, J.M.; Glodek, A.;

rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.

Science 273, 1058-1073, 1996

A:Authors: Kaine, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.

A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannaschii.

A:Reference number: A64300; MUID:96337999; PMID:8688087

A:Accession: A64465

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1005 <BL>

A:Cross-references: GB:U67572; GB:L77117; NID:gl591958; PIDN:AAB99331.1; PID:gl591962; TJ

C:Genetics:

A:Map position: REV1273394-1270377

C:Superfamily: hypothetical protein MJ1322

Query Match 7.5%; Score 179; DB 2; Length 1005;

Best Local Similarity 22.3%; Pred. No. 0.26;

Matches 116; Conservative 80; Mismatches 164; Indels 160; Gaps 27;

Qy 52 DKIKTKYPPENKPGQSNYSFVDNLNLLKA-ITEKEKEKERQSRSSPLDNKLVNEDVD 109

Db 159 DEFEKCY--QKGEIVKEKLERIEGELNKENYEK-----LQNKMSQLE-- 205

Qy 110 STNKRKLIDDYDSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFD 169

Db 206 --KMKLMBINDKLNK- IKKKEPDIETKLPNENKLLYKFKLNK-----EERKALEL 257
QY 170 KIVS-KLNLGLITSOAH-----TLEDEVAEVQLKLSKEANNVEEDPNK 214
Db 258 KNQELKILEYDLNTVVEARETNRHKDEYKYSLVDEIRKIESRL--RELKSHYEDYLK 315
QY 215 PISWTENOAGKIPEKVTWMAAIQDGLAGKENDETVSTLTLTNGLERRTKIYSNDPEEL 274
Db 316 LTKQLEIIKGGDI-EK-----LKEFINKSKYRDDIDNLDTLNKKI-----KDEIERV 360
QY 275 QYFPNYPYALLKSIDSEKAEK-----ETLITIMKTLDFVMMVKYGTISPE 322
Db 361 ETIKOLLEBELKVLNLEIEIKYKRICBECKEYVEKYLESEKAVENKLTLEVITL-- 417
QY 323 EGVSYLENDEMIALQTNKLEKNATD--NISKLPAPS-----EKSHEEDTSTK-- 370
Db 418 -----LQEKSKTEKNINDLETRINKLKEKNIDIESIENSUKIEEEKKV 463
QY 371 -----EEAKMEKEVGLSKDSTK-----DONSPPGK-----TDEPK 405
Db 464 LENLQEKIELNKLGEINSEIKLKLIDELKEVEGCKPLCKTPIDENKQWELNQHT 523
QY 406 E--AYLE-----AIRKNIEWLKKH-----DKGNKEDYDL----- 433
Db 524 QLNKKYTELEINKKIREIKDIEKLBKIDKEENLKLTKLYLEKQSQIEELEKLKNY 583
QY 434 SKMRDFINKQADAYVEKG-----ILD--KEBAEAIKRIYS 466
Db 584 KEQLDEINKKISNVYNGKPVDEILEDIKSLQNLKFKNFYN 623

RESULT 11
B71603
RESA-H3 antigen PFB0915w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C>Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: B71603
R:Gardner, M.J.; Fettel, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: B71603
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1558 <GAR>
A:Cross-references: GB:AE001424; GB:AE001362; NID:93845307; PIDN:AAC71972.1; PID:9384530
A:Experimental source: clone 307
C:Genetics:
A:Gene: PFB0915w

Query Match 7.5%; Score 179; DB 2; Length 1558;
Best Local Similarity 21.5%; Pred. No. 0.45;
Matches 118; Conservative 96; Mismatches 197; Indels 138; Gaps 25;
QY 27 QDKSLHRELAE--RPLNQIAEAEEDKIKKYPENKPGQN-----YSFDNL-- 75
Db 725 EENAVSENVAENLEKLNKTNVTNVLVDKVEETVEISGESLENEMDKAFFSEIFDNVKG 784
QY 76 ---NLK-----AITEKEKIEKERSQIRSSPLDNKLN-----VEDVDSTK-- 112
Db 785 IQENLLTGMFRSITSTVIOSEKVDL--NENVVSSILDNIENKGLLNKLENIISTEGV 843
QY 113 -----NRKLLDDVD-----STKSGLDHFKQDDPGLHQDGTPLTAEDI- 151
Db 844 QETVTEHVQNYYVDVDPAMKQDQFGLINEAGGLKEMFNLED--VFKSESVDITVEBK 902
QY 152 -----VHKIARIYEENDRAVEDKIVSKLNLGLITESTQAHTLEDEVAEVQLKLSKE 204
Db 903 DEPQVEKEKETSIIIEENNIIVD-----VLEEEKEDLTDKMDAVESIEIS 951
QY 205 ANYEEDPNKPTSWTENQAGKIPEKVTWMAAIQDGLAGKENDETVSTLTLTNGLERRTK 264

Db 952 SDSKEE-----TES-----IKQEKDVSLLVEEVQDNDMDSEVKVLELKNMBEELMK 999
QY 265 TYSEDN-----FEELQYFPNFYA--LLKSIDSEKAEK-----EKETLITIMK 304
Db 1000 DAVEINDITSKLIEETQELNEVEADLIKMEKLEKALSSEDSKEIIOAKDDTLEKVI 1059
QY 305 -----TLIDFVMMVKYGTISPE--EGVSYLENDEMIALQTK--NKLEKNATDNISK 354
Db 1060 EEHDITTTLDEV---VELKDVEEDKIEKVSIDLKQLEEDILKEVKEIKELSEIILEDYKEL 1116
QY 355 FPAPSEKSHEDSTKEBAKMEKEVGLSKDSTKDNNSNPGKTDPEKGTAYLEAIRK 414
Db 1117 KITETDILKEKEIEKDHFEKFEBAEBEIKOLADILKEVSVLESVEEEKKLEVEVHE-LKE 1175
QY 415 NIEWLKHHDK--KGNKEDYDLKMRDFINKQADAY---VEKGILDK-----EBA 458
Db 1176 EVEHITISGDAHIKGLEED--DLEEVDDLKGSILDMKGMELGDMDKESLEDVTAKLGERV 1234
QY 459 EAIKRIYS 467
Db 1235 ESLKQVLS 1243

RESULT 12
T28677
rhostry protein - Plasmodium yoelii
C:Species: Plasmodium yoelii
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28677; C45521
R:Keen, J.; Sinha, K.; Brown, K.; Holder, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A:Title: A gene coding for a high molecular mass rhostry protein of Plasmodium yoelii.
A:Reference number: Z20508; MUID:95021522; PMID:7935623
A:Accession: T28677
A:Status: preliminary; translated from GB/EMBL/DDBJ
A:Molecule type: DNA
A:Residues: 1-2269 <KEE>
A:Cross-references: EMBL:127838; NID:9457145; PID:9457146; PIDN:AAA21304.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhostry protein. Multiple co
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: C45521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 2131-2269 <KE2>
A:Cross-references: GB:M34283

Query Match 7.5%; Score 179; DB 2; Length 2269;
Best Local Similarity 21.1%; Pred. No. 0.73;
Matches 113; Conservative 85; Mismatches 193; Indels 144; Gaps 23;
QY 13 LVLIQIAPPKGGQDKSLHRELAEERPLNEQIAEAEEDKIKKYPENKPGQSNYSFV 72
Db 7 LILPLMNQFGLNESMIKLNKSGILRKYTISNQI---KNKLVNSTYPEGREG-----FT 57
QY 73 DNLNL-----LKAITEKEK-----TEKERSIRSSPLDN-----KLN- 104
Db 58 SSLELAKSWEKTKLETITELTKSNEETVRLEKEIREULFKYLDSEAEERKYLEGLKLELNK 117
QY 105 -----VEDVOST-----KNRKLIDD-----YDSTKSGLDH 129
Db 118 KIKDIIAKIEVKNVTLBELKEIKKNAYIDELANQSPYKVTGYIENKNTIYNTIKSYFDQ 177
QY 130 KFQDDPGLHQLDGTPLTAEDIVHKAARIYEENDRAVFKIVSKLNLGLIT-ESQHT 188
Db 178 IYEGDIDTFYN-ELUSSIVKEDPIDDEDTKLENLSKIDNVYDKIQKMBIETVKSHLN 236
QY 189 LE-----DEVAEVLQKLSKEANNVEEDPNKPTSWTENQAGKIPEKVT 231
Db 237 IETNNKLPNTILEKKIYVIDEISKELNKMLEDPFNKEKELSNKISDY-DKKRQLSEYKS 295


```

QY 232 PMAAIQDGLAKGENDT-VSNLTILTNGLERTKTYSEDNFEELQYFPNFVALLKSIDSE 290
Db 296 KMLEIRNY-----NSQTVNDTK-----EAKQNYDKSN-EHMTTIPITNEDEISKIIE 345
QY 291 KEAKEKETLITIMKTLPDVKMMVKYGTISPPEGVSYLENLDEMIALOTKMKLEKNATDN 350
Db 346 VKTM-KDILSKVNTYIDFNK---KYKTVNSEHSQFTLTDKIKAEVSDKELKK----- 396
QY 351 ISKLFPAPPSHEETDS-TKEAAKMEKYGSLKSDKDNPNPGKGTDPKPKTEAYL 409
Db 397 -----CEQSFNDNKLINETKNSIEKEY-----QNINTLKKVDE-----YI 432
QY 410 EAIRKNIEWLKHKDKGKNEDYDLSKMRDPIFKQADAYVEKGILDKEAEAIKRI 464
Db 433 KYVKSTKESITKFSK-----QTILKMLNQNIKTVKETSIDKSYIEKFEQI 480

RESULT 13
A43336
microtubule-vesicle linker CLIP-170 - human
C:Species: Homo sapiens (man)
C:Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C:Accession: A43336
R:Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
Cell 70, 887-900, 1992
A:Title: CLIP-170 links endocytic vesicles to microtubules.
A:Reference number: A43336; MUID:92405160; PMID:1356075
A:Accession: A43336
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1392 <PE>
A:Cross-references: GB:M97501; NID:g180621; PID:AAA35693.1; PID:g180622

Query Match 7.4%; Score 178.5; DB 2; Length 1392;
Best Local Similarity 22.4%; Pred. No. 0.42;
Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;

QY 34 RELSAB-----RPLNEQIAEAEEDKIKKTYPPENKPGQSNVSVFDNL-----NLKA 80
Db 734 RKASSEKSEMKLRQQL-EAAEQIKHLEIEKNAESSKASSITRELQGRELKLNTLQEN 792
QY 81 ITE-----KEKIEKERO-----SIRSPLD--NKL-----NVEDVDSTKN 113
Db 793 LSEVSQVKETLEKELQILKEKFAEASEAVSVQSMQETVNLKHQKEQFNMLSSDLEKL 852
QY 114 RKLIDDYDSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKAARIYEENDRAVFDKIVS 173
Db 853 RENLAD-----MEAKFREKDEREEQL-----IKAKEKLEND-----IA 885
QY 174 KLLNLGLITESQAHTLEDEV-----AEVLQKLISKEANNYE-----EDPNKPTSWTEN 221
Db 886 EIMKMGDNGSSQLTKMDELRLKERDVEELQKLTKANENASFLQKSIDMTVKAESQSQ 945
QY 222 QAGKIPEKVTMAAIQDGLAKGENDT-VSNLTILTNGLERTKTYSEDNFEELQYFPNFY 281
Db 946 EAAKKHEE-----EKKELERKLSL-----LEKMET-SHNQCOELK----- 980
QY 282 ALLKSIDSEKAKEKETLITIMKTLPDVKMMVKYGTISPPEGVSYLENLDEMIALOTKN 341
Db 981 ARYERATSETKHEEILQNLQKLTLLDTEKLG-----ARENSGLLQLEELSRKQAEKA 1036
QY 342 KLEKNATDNISKLFPAPSEKSH--EETDSTKEAAKMEKYGSLKSDKDNPNPGKGT 399
Db 1037 KAAQTAEDAMQIMEQMTKEKTETLASLEDTKQTNKQLNLDLTLENNL-KNVEELNKS 1095
QY 400 EPKGKTEAYLEAIRKNIEWLKHH-----DKGNKEDYDLSKMRDPIFKQADAYV 448
Db 1096 ELLTVENQKMEEFREKTEITLKQAAQKSQQLSALQENNVKLAEELGRSDRDEVTSHQLEE 1155
QY 449 EKGILDKEAEAIKRI 463
Db 1156 ERSVLNNQLLEMKKR 1170

```

```

RESULT 14
S22695
restin - human
C:Species: Homo sapiens (man)
C:Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Nov-1999
C:Accession: S22695; S19853
R:Balbe, G.; Delabie, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;
EMBO J. 11, 2103-2113, 1992
A:Title: Restin: a novel intermediate filament-associated protein highly expressed in the
A:Reference number: S22695; MUID:92289675; PMID:1600942
A:Accession: S22695
A:Molecule type: mRNA
A:Residues: 1-1427 <BI>
A:Cross-references: EMBL:X64838; NID:g35998; PIDN:CAA46050.1; PID:g35999
C:Keywords: cytoskeleton

Query Match 7.4%; Score 178.5; DB 2; Length 1427;
Best Local Similarity 22.4%; Pred. No. 0.43;
Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;

QY 34 RELSAB-----RPLNEQIAEAEEDKIKKTYPPENKPGQSNVSVFDNL-----NLKA 80
Db 769 RKASSEKSEMKLRQQL-EAAEQIKHLEIEKNAESSKASSITRELQGRELKLNTLQEN 827
QY 81 ITE-----KEKIEKERO-----SIRSPLD--NKL-----NVEDVDSTKN 113
Db 828 LSEVSQVKETLEKELQILKEKFAEASEAVSVQSMQETVNLKHQKEQFNMLSSDLEKL 887
QY 114 RKLIDDYDSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKAARIYEENDRAVFDKIVS 173
Db 888 RENLAD-----MEAKFREKDEREEQL-----IKAKEKLEND-----IA 920
QY 174 KLLNLGLITESQAHTLEDEV-----AEVLQKLISKEANNYE-----EDPNKPTSWTEN 221
Db 921 EIMKMGDNGSSQLTKMDELRLKERDVEELQKLTKANENASFLQKSIDMTVKAESQSQ 980
QY 222 QAGKIPEKVTMAAIQDGLAKGENDT-VSNLTILTNGLERTKTYSEDNFEELQYFPNFY 281
Db 981 EAAKKHEE-----EKKELERKLSL-----LEKMET-SHNQCOELK----- 1015
QY 282 ALLKSIDSEKAKEKETLITIMKTLPDVKMMVKYGTISPPEGVSYLENLDEMIALOTKN 341
Db 1016 ARYERATSETKHEEILQNLQKLTLLDTEKLG-----ARENSGLLQLEELSRKQADKA 1071
QY 342 KLEKNATDNISKLFPAPSEKSH--EETDSTKEAAKMEKYGSLKSDKDNPNPGKGT 399
Db 1072 KAAQTAEDAMQIMEQMTKEKTETLASLEDTKQTNKQLNLDLTLENNL-KNVEELNKS 1130
QY 400 EPKGKTEAYLEAIRKNIEWLKHH-----DKGNKEDYDLSKMRDPIFKQADAYV 448
Db 1131 ELLTVENQKMEEFREKTEITLKQAAQKSQQLSALQENNVKLAEELGRSDRDEVTSHQLEE 1190
QY 449 EKGILDKEAEAIKRI 463
Db 1191 ERSVLNNQLLEMKKR 1205

RESULT 15
D96796
probable heat shock protein, 53413-59028 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: D96796
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;
ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

```

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A;Reference number: A86141; MUID:21016719; PMID:11130712
A;Accession: D96796
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-1871 <STO>
A;Cross-references: GB:AE005173; NID:G6143906; PIDN:AAF0452.1; GSPDB:GN00141
C;Genetics:
A;Gene: F28O16.15
A;Map position: 1

```
Query Match          7.4%; Score 177; DB 2; Length 1871;
Best Local Similarity 20.7%; Pred. No. 0.71;
Matches 116; Conservative 85; Mismatches 212; Indels 148; Gaps 22;

QY 12 VLVLPQAFPKP--GGSQDKSLHNRLSALRPLNEQIAAEEDKIKTYPPENKPGOSNY 69
Db 1052 VLVEEETPKKHGTGGEHDNDHKEEQENVIKAEALNTEEDSPKVEIEKQ----- 1105

QY 70 SFVDNLNLKAI TEKEIEKERQSI RSPLOKLNKLVEDVDSTKN--RKLDDYDSTKSG 126
Db 1106 ---DHGELKRSVMQAKRQSTEEKD-KTRAMEKNETVERRKQTKDGLSLKLGREGDEPGLG 1161

QY 127 LHKFQDDPDGLHQLDGTPLT--AEDIVHKTAARIYEENDRAVFD-----KIVSKLLNLGL 180
Db 1162 --HERRGEEDRIEELVETEISDHKEKVKKQDEYILRSQDTGKVDLGERERRSKQRKIHK 1219

QY 181 ITESQAHTLEDEVAEVLQKLISKEAN-----NVEEDPNKPTSWTENQAGKIPEKVTP- 232
Db 1220 SVEDEIGQOEDAEAAAVSRNENGSRKVQTIIESEK-----HKEQNKIPETSNPE 1274

QY 233 -----MAAIQDGLAKGEN-----DETVSNTLTLTNGLERRTKT 265
Db 1275 VNEEDEERVVEKETKEVEAHVQELGKTENCKDDDGGERREERCKQGTAEENMLRQRFKT 1334

QY 266 YSEDNFEELQYFPNFYALLKSID--SEKEAKEKTLITIMKTLIDFVMVMKYGTI----- 319
Db 1335 KSDD-----GIVRKIQETKEEPEDEKKS---QESSSHVVVKLVAEEDGSLRNL 1378

QY 320 ---SPEGVSYLENLDEMIALQTKNLEKNATD--NISKLFPAPSEKSHSEETDST----- 369
Db 1379 EFSEKESTVSKMLIDESKEEHEHKIRKPTTEERSNAPVIEKQGNKKNAAEENQDKIDRR 1438

QY 370 -KEEAAMKEKEYGSLK-----STKDDNSNPGKTD-- 399
Db 1439 GKQEIKGQEPYGLVRNGEHDKI TEYHRGEEKGTAEENVSTKIQTQKDELEKPKPSEIS 1498

QY 400 -----EPKGKTEAYLEAIRKNIEMLKKHDKKGNKEDYDLS-KMRDFIN 441
Db 1499 ENHNIHEFMDSSQSDIEEKGSQDAEKYAKQNKIQEVMNDEDK--KEEYHISERVNEMA 1556

QY 442 K---QADAYVEKGILDKERAE 459
Db 1557 KRILQVESKANDGSSKKNETE 1577
```

Search completed: January 2, 2004, 12:36:54
Job time : 22 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 12:33:21 ; Search time 18 Seconds

(without alignments)
1222.694 Million cell updates/sec

Title: ABU66614

Perfect score: 2399

Sequence: 1 MGFLGTGTVLVLVLPQAF.....EKGILDKBEAEAKRIYSSL 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2395	99.8	468	1	SG3_HUMAN	Q8wxd2 homo sapien
2	2102.5	87.6	471	1	SG3_MOUSE	P47867 mus musculus
3	2096.5	87.4	471	1	SG3_RAT	P47868 rattus norv
4	191	8.0	2464	1	MAPB_MOUSE	P14873 mus musculus
5	186.5	7.8	2116	1	MY52_DICDI	P08799 dictyosteli
6	183.5	7.6	1790	1	USO1_YEAST	P25386 saccharomyc
7	180.5	7.5	2459	1	MAPB_RAT	P15205 rattus norv
8	179	7.5	1005	1	RA50_METJA	Q58718 methanococc
9	178.5	7.4	1427	1	REST_HUMAN	P30622 homo sapien
10	178	7.4	882	1	RA50_PYRFU	P59301 pyrococcus
11	172.5	7.2	1164	1	BAG_STRAG	P27951 streptococc
12	169	7.0	700	1	TRDN_CANFA	P82179 canis famil
13	167.5	7.0	957	1	KP5C_HUMAN	O60282 homo sapien
14	166.5	6.9	1875	1	MU1P_YEAST	Q2455 saccharomyc
15	166	6.9	2663	1	CENE_HUMAN	Q02224 homo sapien
16	166	6.9	5038	1	PCLO_MOUSE	Q9qyx7 mus musculus
17	163.5	6.8	956	1	KP5C_MOUSE	P28738 mus musculus
18	163	6.8	1251	1	RBP2_PLAYB	Q00799 plasmodium
19	161.5	6.7	1360	1	CING_XENLA	Q9p7d7 xenopus lae
20	161.5	6.7	1969	1	MYSA_CAEEL	P22844 caenorhabdi
21	161	6.7	852	1	RA50_THEMEA	Q9xlx1 thermotoga
22	159.5	6.6	997	1	CPH1_RAT	Q03410 rattus norv
23	159.5	6.6	1978	1	MYHB_CHICK	P10587 gallus gall
24	159	6.6	1208	1	PCP1_SCHPO	Q92351 schizosacch
25	158.5	6.6	1526	1	MY52_SCHPO	Q9usi6 schizosacch
26	158	6.6	1972	1	MYHB_RABIT	P35748 cryptolagus
27	157.5	6.6	1972	1	MYHB_HUMAN	P35749 homo sapien
28	157.5	6.6	2230	1	GOG4_HUMAN	Q13439 homo sapien
29	157	6.5	727	1	MPF1_ARATH	Q91w85 arabidopsis
30	157	6.5	944	1	NUF1_YEAST	P32380 saccharomyc
31	157	6.5	2017	1	MY5N_DRONE	Q93323 drosophila
32	156.5	6.5	1087	1	AKA9_RABIT	Q28628 cryptolagus
33	156.5	6.5	2468	1	MAPB_HUMAN	P46821 homo sapien

34	156	6.5	705	1	TRDN_RABIT	Q28820 oryctolagus
35	155.5	6.5	1130	1	YLI1_CAEEL	Q11102 caenorhabdi
36	155.5	6.5	5085	1	PCLO_RAT	Q9jks6 rattus norv
37	154.5	6.4	1539	1	Y373_HUMAN	O15078 homo sapien
38	154	6.4	1433	1	REST_CHICK	O42184 gallus gall
39	154	6.4	1957	1	SPOF_SCHPO	O10411 schizosacch
40	154	6.4	1962	1	MYSA_DRONE	P05661 drosophila
41	153.5	6.4	1976	1	MYHA_RAT	Q9jlco rattus norv
42	153.5	6.4	2104	1	MY53_SCHPO	O14157 schizosacch
43	152.5	6.4	2022	1	ANTI_ONCOC	P21249 onchocerca
44	152.5	6.4	8545	1	ANCI_CAEEL	Q9n4m4 caenorhabdi
45	152	6.3	667	1	CYLI_BOVIN	P35662 bos taurus

ALIGNMENTS

RESULT 1
SG3_HUMAN STANDARD; PRT; 468 AA.
AC Q8WXD2; Q96C83; Q96GE8; Q9Y6G7;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Secretogranin III precursor (SgIII).
GN Sg3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pancreatic;
RA Song H., Peng Y., Huang Q., Dai M., Mao Y., Zhang Q., Mao M., Fu G.,
RA Luo M., Chen J., Hu R.;
RT "Human secretogranin III mRNA, complete cds";
RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RX MEDLINE=22093564; PubMed=12098761;
RA Rong Y.P., Liu F., Zeng L.C., Ma W.J., Wei D.Z., Han Z.G.;
RT Cloning and characterization of a novel human secretory protein:
RT secretogranin III.;
RA Acta Biochim. Biophys. Sin. 34:411-417(2002).
[3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain, and Lung;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalls D.E.,
RA Schnerch A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- SUBCELLULAR LOCATION: Secreted. Neuroendocrine and endocrine
CC secretory granules.
CC -!- TISSUE SPECIFICITY: Expressed in brain, heart, kidney, liver and
CC skeletal muscle.

OS Rattus norvegicus (Rat).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 OX NCBI_TaxID=10116;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=95001263; PubMed=7917832;
 RA Dopazo A., Lovenberg T.W., Danielson P.E., Ottiger H.-P.,
 RA Sutcliffe J.G.;
 RT "Primary structure of mouse secretogranin III and its absence from
 RT deficient mice.";
 RL J. Mol. Neurosci. 4:225-233(1993).
 RN [2]
 RP PRELIMINARY SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=90376160; PubMed=2204688;
 RA Ottiger H.-P., Battenberg E.F., Tsou A.-P., Bloom F.E.,
 RA Sutcliffe J.G.;
 RT "18075: a brain- and pituitary-specific mRNA that encodes a novel
 RT chromogranin / secretogranin-like component of intracellular
 RT vesicles.";
 RL J. Neurosci. 10:3135-3147(1990).
 CC -1- SUBCELLULAR LOCATION: Secreted. Neuroendocrine and endocrine
 CC secretory granules.
 CC -1- TISSUE SPECIFICITY: BRAIN- AND PITUITARY-SPECIFIC.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; U02983; AA56637.1; -;
 KW Signal; Cleavage on pair of basic residues.
 FT SIGNAL 1 22 POTENTIAL.
 FT CHAIN 23 471 SECRETOGRANIN III.
 SQ SEQUENCE 471 AA; 53183 MW; 977B3F8885F33223 CRC64;
 Query Match 87.4%; Score 2096.5; DB 1; Length 471;
 Best Local Similarity 87.9%; Pred. No. 1.1e-93;
 Matches 414; Conservative 18; Mismatches 36; Indels 3; Gaps 1;
 QY 1 MGFLGTGWILVLVL---PIQAFKPGSQSKLHNRLSALPNEQIAEADKIKKT 57
 DB 1 MGFLTGSWILVLVNSGPIQAFKPGSQSKLHNRLSALPNEQIAEADKIKKT 60
 QY 58 YPPENKQSQSYFVDNINLLKAITEKEKIEKRSQSRSSPLDNKLVNVDVSTKNRKL 117
 DB 61 YPSESQPSSESNFSDNINLLKAITEKETVEKAKQSRSSPFDNRLNVDDADSTKNRKL 120
 QY 118 DDYSTKSLDHHKQDDPDGLHOLDGTLTAEDIVHKIARIYEENDRAVDFKIVSKLLN 177
 DB 121 DEYDSTKSLGRKVDQDDPDGLHOLDGTLTAEDIVHKIARIYEENDRGVDFKIVSKLLN 180
 QY 178 LGLITESQAHTLEDAVAVLOKLSKEANNYEEDNKPTSWENQAGKIPKVPMAIQ 237
 DB 181 LGLITESQAHTLEDAVAVLOKLSKEANNYEEDNKPTSWENQAGKIPKVPMAIQ 240
 QY 238 DGLAKGENDETVSNLTITNGLRRTKTSYEDNFEELQYFNFVALLKSIDSEKEAKE 297
 DB 241 DGFNTRENDVTSNLTITNGLRRTNPHRDDDFEELQYFNFVALLTSIDSEKEAKE 300
 QY 298 TLITIMKTLIDFVKMWYKGTISPEEGVSYLENDEMIALQTKNLEKNTDNTSKLPPA 357
 DB 301 TLITIMKTLIDFVKMWYKGTISPEEGVSYLENDEMIALQTKNLEKNTDNTSKLPPA 360
 QY 358 PSEKSHETDSTKEAAKWEKEYGSLKSDTDDNSNPGKTDPEKGTAEVLEAIRKNIE 417
 DB 361 PPEKSHETDSTKEAAKWEKEYGSLKSDTDDNSNPGKTDPEKGTAEVLEAIRKNIE 420
 QY 418 WLKXHKDKGNKEDYDLSKWRDFINKQADYVVEKGLDKKEAEAIKRIYSSL 468

Db 421 WLKXHKDKGNKEDYDLSKWRDFINKQADYVVEKGLDKKEAEAIKRIYSSL 471
 RESULT 4
 MAPB_MOUSE
 ID MAPB_MOUSE STANDARD; PRT; 2464 AA.
 AC P14873;
 DT 01-APR-1990 (Rel. 14, Created)
 DT 01-APR-1990 (Rel. 14, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Microtubule-associated protein 1B (MAP1B) (MAP1.2) (MAP1(X))
 DE [Contains: MAP1 light chain LC1].
 GN MAP1B OR MTAP1B OR MTAP5.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A., AND DOMAIN.
 RC STRAIN=Swiss Webster; TISSUE=Brain;
 RX MEDLINE=90094539; PubMed=2480963;
 RA Noble M., Lewis S.A., Cowan N.J.;
 RT "The microtubule binding domain of microtubule-associated protein
 RT MAP1B contains a repeated sequence motif unrelated to that of MAP2
 RT and tau.";
 RL J. Cell Biol. 109:3367-3376(1989).
 CC -1- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
 CC PHOSPHORYLATED MAP1B MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
 CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAP1B BINDS TO AT LEAST
 CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
 CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
 CC STABILIZING MICROTUBULES.
 CC -1- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
 CC WITH MAP1A AND MAP1B PROTEINS.
 CC -1- DOMAIN: Has a highly basic region with many copies of the sequence
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
 CC responsible for the binding of MAP1B to microtubules.
 CC -1- PTM: LC1 IS COEXRESSED WITH MAP1B. IT IS A POLYPEPTIDE GENERATED
 CC FROM MAP1A BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
 CC BOTH MAP1A AND MAP1B. IT INTERACTS WITH THE AMINO-TERMINAL REGION
 CC OF MAP1B.
 CC -1- SIMILARITY: TO MAP1A.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC
 DR EMBL; X51396; CAA35761.1; -;
 DR PIR; S07549; QRMSP1.
 DR MGD; MGI:1306778; Mtap1b.
 DR GO; GO:0016358; P:dendrite morphogenesis; IMP.
 DR GO; GO:0001578; P:microtubule bundling; IMP.
 DR InterPro; IPR00102; MAP1b_neuraxin.
 DR Pfam; PF00414; MAP1b_neuraxin; 10.
 DR PROSITE; PS00230; MAP1B_NEURAXIN; 7.
 KW Microtubules; Repeat; Phosphorylation.
 FT CHAIN ? 2464 MAP1 LIGHT CHAIN LC1.
 FT REPEAT 1874 1890 MAP1B 1.
 FT REPEAT 1891 1907 MAP1B 2.
 FT REPEAT 1908 1924 MAP1B 3.
 FT REPEAT 1925 1941 MAP1B 4.
 FT REPEAT 1942 1958 MAP1B 5.
 FT REPEAT 1959 1975 MAP1B 6.
 FT REPEAT 1993 2009 MAP1B 7.
 FT REPEAT 2010 2026 MAP1B 8.
 FT REPEAT 2027 2043 MAP1B 9.
 FT REPEAT 2044 2060 MAP1B 10.
 FT DOMAIN 589 787 LYS-RICH (HIGHLY BASIC, CONTAINS MANY

```

FT  SQ  SEQUENCE  2464 AA;  270408 MW;  FBD3DD99CFBDA87 CRC64;
      Query Match      8.0%;  Score 191;  DB 1;  Length 2464;
      Best Local Similarity 21.3%;  Pred.No.0.098;
      Matches 118;  Conservative 85;  Mismatches 186;  Indels 166;  Gaps 24;

QY  22  KPGSQDKSLHNRELSAERPLNEQIAEA-----EEDKIKKTYP--PENKPGQSNYSFVD  73
DQ  550  KVASKSVRKESKEETPEVTKTSQVEKTPKVESKEKVLVKDKPKVTKESKP-----  600
QY  74  NLNLLKATEKIEKIEKESQSTRSSPLNKLNVEDVDSTGNKRLIDVDYDSTKGLDHRFQD  133
DQ  601  -----SVTEKEVSKESQSPVKAEBQAEQKQTESKPKVTQKVKVKEIKTK--LEEKKEE  652
QY  134  DP--DGLHQLDQTP--TAEDIVHKIAARIYEENDRAVDFKIVKISKLNLGLIETESQ  185
DQ  653  KPKKEVVKEDKTPDKDEKPKKEVKEIKKEIKKE--ERKELKKEVKK-----ETP  703
QY  186  AHTLEDEVAEVLQKLISKEANNYEDPNK-----PTSWTENQAGKIPE---  228
DQ  704  LKDAKKEVKEKKEVKEK-----EKEPKKEIKKIDKIKKSTPQSDTKKPSALKPKVAKK  759
QY  229  ----KVTPMAA--TQD-CLAK-----GENDETVENTL-----  253
DQ  760  BESTKKEPLAAGKLDKGVKVIKKEGKTTAAATAVGTAATTAAVVAAAGIAASGPVKE  819
QY  254  ----TITNGLERTKTYSEDNFEELOVFPNFYALLKSIDSEKAKE-----KET  298
DQ  820  LEAERSLWSSPEDLTQPEELKAEIDVAKDIKQLELIEDDEKKEKTKQGEAVVIQKET  879
QY  299  LITMTKLIDFVKMVKVGTITSPREGVSYLENLDEMIALQTKNKL-----ERNADTNISKL-  354
DQ  880  EVS-----KGSABSPDEGITTTEGECE--CEQTPELEPVEKQGVDDIEKPE  924
QY  355  -----EPAPSEKSHEDTSTEEAAKKEVGSGLKSDTKDNDNPNPGKTDPEKPKTEAYL  409
DQ  925  DEGAGFSESSTGTYEKEATEEAEPEEP--GEDNASGSASKHSPTDEDDSAKAEADVHL  983
QY  410  EAIRKNI-----EWLKK-----HDKGNKNKEDY-----DLSKMRDFIN  441
DQ  984  KKRSESVVSGDRAEEDMDVLEKGEAEQSEEGEEDKAEADAREGEYFDPKTEAEDVYM  1043
QY  442  KQADAYVEKGLDKKE  456
DQ  1044  AVADKAAEAGVTEQ  1058

RESULT 5
ID  MYS2  DICDI  STANDARD;  PRT;  2116 AA.
AC  P08799;
DT  01-NOV-1988 (Rel. 09, Created)
DT  01-OCT-1989 (Rel. 12, Last sequence update)
DT  15-SEP-2003 (Rel. 42, Last annotation update)
DE  Myosin II heavy chain, non muscle.
GN  MhCA.
OS  Dictyostelium discoideum (Slime mold).
OC  Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
OX  NCBI_TaxID=44689;
RN  [1]
RP  SEQUENCE FROM N.A.
RX  MEDLINE=87092266; PubMed=3540939;
RT  Warrick H.M., de Lozane A., Leinwand L.A., Spudich J.A.;
RA  "Conserved protein domains in a myosin heavy chain gene from
RT  Dictyostelium discoideum."
RL  Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
RN  [2]
RP  PHOSPHORYLATION SITES, AND MUTAGENESIS.
RX  STRAIN=AX2;
RL  MEDLINE=90353583; PubMed=2387408;
RA  Lueck-Vielmeier D., Schleicher M., Grabatin B., Wippler J.,
      Gerisch G.;

```

```

RT  "Replacement of threonine residues by serine and alanine in a
RL  phosphorylatable heavy chain fragment of Dictyostelium myosin II."
RN  FEBS Lett. 269:239-243(1990).
RP  [3]
RX  PHOSPHORYLATION SITES.
RL  MEDLINE=88112226; PubMed=2828113;
RX  Wagle G., Noegel A., Scheel J., Gerisch G.;
RA  "Phosphorylation of threonine residues on cloned fragments of the
RT  Dictyostelium myosin heavy chain."
RL  FEBS Lett. 227:71-75(1988).
RN  [4]
RX  X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RL  MEDLINE=95345066; PubMed=7619795;
RX  Fisher A.J., Smith C.A., Thoden J.B., Smith R., Holden H.M.,
RA  Rayment I.;
RL  "X-ray structures of the myosin motor domain of Dictyostelium
RT  discoideum complexed with MgADP.Befx and MgADP.ALf4-."
RL  Biochemistry 34:8960-8972(1995).
RN  [5]
RX  X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RL  MEDLINE=95345067; PubMed=7619796;
RX  Smith C.A., Rayment I.;
RA  "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RT  truncated head of Dictyostelium discoideum myosin to 2.7-A
RL  resolution."
RL  Biochemistry 34:8973-8981(1995).
RN  [6]
RX  X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RL  MEDLINE=96206189; PubMed=8611530;
RX  Smith C.A., Rayment I.;
RT  Dictyostelium discoideum myosin motor domain to 1.9-A resolution."
RL  Biochemistry 35:5404-5417(1996).
RN  [7]
RX  X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RL  MEDLINE=97452580; PubMed=9305951;
RX  Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RA  "X-ray structures of the MgADP. MgATPgamas, and MgAMPPNP complexes
RT  of the Dictyostelium discoideum myosin motor domain."
RL  Biochemistry 36:11619-11628(1997).
RN  [8]
RX  X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RL  MEDLINE=98070605; PubMed=9405148;
RX  Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RA  "X-ray crystal structure and solution fluorescence characterization
RT  of Mg.2'(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RL  Dictyostelium discoideum myosin motor domain."
RL  J. Mol. Biol. 274:394-407(1997).
CC  -!- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC  ACTIVITY THAT IS ACTIVATED BY ACTIN.
CC  -!- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC  INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC  LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC  (MLC-2).
CC  -!- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC  CORTEX.
CC  -!- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC  MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
CC  SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC  SUBFRAGMENT (S2).
CC  -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, SHOWING
CC  CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC  CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC  -!- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC  THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC  -!- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC  ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC  POSITION (688).
CC  -!- SIMILARITY: Contains 1 myosin-like globular head domain.
CC  -!- SIMILARITY: Contains 1 IQ domain.
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -

```



```

RESULT 6
USOI_YEAST
ID USOI_YEAST STANDARD; PRT; 1790 AA.
AC P25386;
DT 01-MAY-1992 (Rel. 22, Created)
DT 01-MAY-1992 (Rel. 22, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Intracellular protein transport protein USOI.
GN USOI OR INT1 OR YDL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=X2180-1A;
RX MEDLINE=91185402; PubMed=2010462;
RA Nakajima H., Hirata A., Ogawa Y., Yonehara T., Yoda K.,
RA Yamasaki M.;
RT "A cytoskeleton-related gene, usol, is required for intracellular
RT protein transport in Saccharomyces cerevisiae.";
RL J. Cell Biol. 113:245-260(1991).
RN [2]
SEQUENCE OF 782-1790 FROM N.A.
RP Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,
RA Kendrick K.E.;
RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE OF 1-8 FROM N.A.
RP Bai Y., Symington L.S.;
RA Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI
CC COMPLEX.
CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR
CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE
CC ER AND THE GOLGI COMPLEX.
CC -!- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED
CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL
CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.
CC -!- SIMILARITY: BELONGS TO THE VDP/USOI/YBL047C FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC -----
CC EMBL; X54378; CAA38253.1; -
CC EMBL; L03188; AAB00143.1; -
CC EMBL; U53668; AAB66659.1; -
CC SGD; S0002216; USOI.
CC InterPro; IPR002017; Spectrin.
CC InterPro; IPR006955; Usol_p115_C.
CC InterPro; IPR006953; Usol_p115_head.
CC Pfam; PF04871; Usol_p115_C; 1.
CC Pfam; PF04869; Usol_p115_head; 1.
CC TRANSPORT; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.
CC DOMAIN 1 724 GLOBULAR HEAD.
CC DOMAIN 725 1790 COILED COIL (POTENTIAL).
CC DOMAIN 465 487 CHARGED (HYPER-HYDROPHILIC).
CC DOMAIN 991 1790 DISPENSABLE FOR THE PROTEIN FUNCTION.
CC DOMAIN 1172 1786 ASP/GLU-RICH (ACIDIC).
CC CONFLICT 847 847 G -> E (IN REF. 2).
CC CONFLICT 924 924 E -> K (IN REF. 2).
CC CONFLICT 1253 1253 V -> I (IN REF. 2).
CC CONFLICT 1319 1319 I -> V (IN REF. 2).
CC CONFLICT 1461 1461 N -> S (IN REF. 2).
CC CONFLICT 1581 1581 G -> S (IN REF. 2).
CC CONFLICT 1600 1600 I -> V (IN REF. 2).
CC CONFLICT 1661 1661 R -> S (IN REF. 2).
CC CONFLICT 1772 1772 D -> DEEDDEE (IN REF. 2).

```

```

SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9PD4818 CRC64;
Query Match 7.6%; Score 183.5; DB 1; Length 1790;
Best Local Similarity 20.3%; Pred. NO. 0.15;
Matches 120; Conservative 91; Mismatches 200; Indels 179; Gaps 23;
QY 29 KSLHN--RELSAERPLNEQIABAEDKIKKTYPPENKPGQSNVFDVNLNLLKAI---T 82
DB 965 KSLANNYKMQAE---NESLIKAVEE-----SKNESSIQLSNLQNKQSDMSQ 1008
QY 83 EKEKIKERQSRSSPLDNKLVNVEDVSTYKNRKLI-----DDYSTKSGLDHKKFQD--- 133
DB 1009 EKENFOIERGSIKKNIEQLKKTISLEQTKBEIISKSDSSKDEYSQISLLKLELTATT 1068
QY 134 -DPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGLITESQAH----- 187
DB 1069 ANDENVNKISIELTKTRELEAEALAA--YKNLKNVELETLETSEKALKEVKEEHLKEEK 1126
QY 188 -TLEDEVAEVLQKLISKAN-----NYEED--PNKPTSWTENQAGKIPEK 229
DB 1127 IQLEKEATETKQOLNSLRANLESLEKEHEDLAAQLKYEEQIANKERQYNE-EISQLNDE 1185
QY 230 VTPMAIQDGLAKGENDETVSNTLTITNGLRRT-----KTYSEDN-- 270
DB 1186 ITSTOQENESIKK-KNDELEGEVKAAMKSTSEQSNLKKSEIDALNLQIKELKKNETNEA 1244
QY 271 -----FEELOYFPNF-----YALLKSIDSEKAK 294
DB 1245 SLESIKSVESVETVKKIKELQDCNFKPEKVESELEDKDKASEDKNSKYLELQK-ESEKIKE 1303
QY 295 EKETLITIMKTILDFVGMVK-----KTSSEERKNAEQLKLNQIKNQ 1363
DB 1304 ELDAKTELKIQLEKITNLSKAKESSELSLKKTSSEERKNAEQLKLNQIKNQ 1363
QY 316 -----YGTISPE--EGVSYLENLDEMIALQTKNKLKKNATONISKLPAPSEK 361
DB 1364 AFERKLLNEGSSITQISEKINTLE--DELIRLQNLQNELKAKEIDNTRSELEKVSLS 1421
QY 362 SHEETDSTKEAAKKEVGVSKDS-TKDDNSNPGKTDPE-----KGKTEAYLEAIRK 414
DB 1422 NDLLEEKQNTKISQDEILSYKDKITRNDKLSIERDKRDLSEKQELRAAQESKAK 1481
QY 415 NIEMLKHKDKKNGEDYDLKMRDFFINQADAYVEKGILD-KEEAFAIKR 463
DB 1482 VEEGLKLEESSKBAELESKEKEMM-KKLESTTIESNETLKSSMETIRK 1530
RESULT 7
MAPB_RAT
ID MAPB_RAT STANDARD; PRT; 2459 AA.
AC P15205; Q62958; Q9ER21; Q9QW92;
DT 01-APR-1990 (Rel. 14, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated protein 1B (MAP 1B) (Neuraxin) [Contains: MAP1
DE light chain LC1].
GN MAP1B.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
SEQUENCE OF 1-142 FROM N.A.
RC STRAIN=Sprague-Dawley; TISSUE=Testis;
RX MEDLINE=96257242; PubMed=8666295;
RA Liu D., Fischer I.;
RT "Isolation and sequencing of the 5' end of the rat microtubule-
RT associated protein (MAP1B)-encoding cDNA.";
RL Gene 172:307-308(1996).
RN [2]
SEQUENCE OF 96-2459 FROM N.A.; DOMAIN, AND INDUCTION.
RC STRAIN=Sprague-Dawley; TISSUE=Brain, and Glial tumor;
RX MEDLINE=92347374; PubMed=1639092;

```


RA Zauner W., Kratz J., Staunton J., Feick P., Wiche G.;
 RT "Identification of two distinct microtubule binding domains on
 RL recombinant rat MAP 1B.";
 RN Eur. J. Cell Biol. 57:66-74 (1992).
 [3]
 RP SEQUENCE OF 1541-2459 FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE-Spinal cord;
 RX MEDLINE=90059871; PubMed=255150;
 RA Rientz A., Grenningloh G., Hermans-Borgmeyer I., Kirsch J.,
 RA Littauer U.Z., Prior P., Gundelfinger E.D., Schmitt B., Betz H.;
 RT "Neuraxin, a novel putative structural protein of the rat central
 RT nervous system that is immunologically related to microtubule-
 RT associated protein 5.";
 RL EMBO J. 8:2879-2888 (1989).
 [4]
 RN DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.
 RP MEDLINE=97405699; PubMed=9260743;
 RX Ma D., Nothias F., Boyne L.J., Fischer I.;
 RA "Differential regulation of microtubule-associated protein 1B (MAP1B)
 RT in rat CNS and PNS during development.";
 RL J. Neurosci. Res. 49:319-332 (1997).
 CC -!- FUNCTION: The function of brain MAPs is essentially unknown.
 CC Phosphorylated MAP1B may play a role in the cytoskeletal changes
 CC that accompany neurite extension. Possibly MAP1B binds to at least
 CC two tubulin subunits in the polymer, and this bridging of subunits
 CC might be involved in nucleating microtubule polymerization and in
 CC stabilizing microtubules.
 CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
 CC with MAP1A and MAP1B proteins.
 CC -!- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem,
 CC cerebellum and cerebrum). Not expressed in liver, spleen, kidney,
 CC heart or muscle.
 CC -!- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic
 CC nerve levels are high early in development but decrease during
 CC postnatal development and are low in adults. In dorsal root
 CC ganglia levels remain high throughout development.
 CC -!- INDUCTION: By nerve growth factor.
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
 CC responsible for the binding of MAP1B to microtubules.
 CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated
 CC from MAP1B by proteolytic processing. It is free to associate with
 CC both MAP1A and MAP1B. It interacts with the amino-terminal region
 CC of MAP1B (By similarity).
 CC -!- PTM: Phosphorylated.
 CC -!- SIMILARITY: TO MAP1A.
 CC -!- CAUTION: A C-terminal fragment of this protein (residues 1597 to
 CC 2459) was originally described as neuraxin in Ref.3.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

 DR EMBL; U52950; AAB17068.1; -;
 DR EMBL; X60370; CAC16162.1; -;
 DR EMBL; X16623; CAA34620.1; ALT_SEQ.
 DR FIR; A56577; A56577.
 DR InterPro; IPR000102; MAP1B neuraxin.
 DR Pfam; PF00414; MAP1B neuraxin; 10.
 DR PROSITE; PS00230; MAP1B NEURAXIN; 8.
 KW Microtubules; Repeat; Phosphorylation.
 FT CHAIN ? 2459
 FT REPEAT 1869 1885 MAP1 LIGHT CHAIN LC1.
 FT REPEAT 1886 1902 MAP1B 1.
 FT REPEAT 1903 1919 MAP1B 2.
 FT REPEAT 1920 1936 MAP1B 3.
 FT REPEAT 1937 1953 MAP1B 4.
 FT REPEAT 1954 1970 MAP1B 5.
 FT REPEAT 1988 2004 MAP1B 6.
 FT REPEAT 1988 2004 MAP1B 7.

FT REPEAT 2021 2005 2021 MAPIB 8.
 FT REPEAT 2022 2038 MAPIB 9.
 FT REPEAT 2039 2055 MAPIB 10.
 FT DOMAIN 559 1035 GLU-RICH.
 FT DOMAIN 588 786 LYS-RICH (HIGHLY BASIC, CONTAINS MANY
 FT KKEE AND KKEI/V REPEATS).
 FT DOMAIN 2224 2312 LYS-RICH.
 FT CONFLICT 127 127 M -> V (IN REF. 1).
 FT CONFLICT 140 140 T -> S (IN REF. 1).
 FT CONFLICT 2112 2112 R -> K (IN REF. 3).
 FT CONFLICT 2169 2169 L -> I (IN REF. 3).
 SQ SEQUENCE 2459 AA; 269497 MW; 2E3F6872DDEB8BA2 CRC64;
 Query Match 7.5%; Score 180.5; DB 1; Length 2459;
 Best Local Similarity 22.7%; Pred. No. 0.31;
 Matches 120; Conservative 77; Mismatches 200; Indels 131; Gaps 22;
 Qy 22 KPGSQDQSLNRLSALRPLNEQIAEAEEDKIK-KYPPENKPGSQSYFVDMNLNLKA 80
 Db 549 KPLSSKSVRKESKEAPEATKASQVETPKVESKEKIVVKKDKPKG-----VESKPS 600
 Qy 81 ITEKEKTEKERSIRSSPLDNKLNVDDVSTKNRKLDDVDSTKSLDGHKFQDDP--DGL 138
 Db 601 VTEKEVSKESQSPVKAFAEAKAATESKPKVKDKVVKKEIKTP--EEKKEEPKKEVA 658
 Qy 139 HQLDGTPLTADIVHKIAARIYEENDRAVFDKIVSKLLNLGLITTESQAH--LEDEVAEVL 197
 Db 659 KKEDKTPLKXDEKPKK-----EAKKEIKKEIKKEKELKKEVKETPLKDAKKEV- 710
 Qy 198 QKLISKANVEEDPNK-----PTSTENQAG-----KIPEKVTPTMAA- 235
 Db 711 KDKKEKVEKEEPKKEIKKIDIKKSTPLSDTKKPAALKPKVAKKEEPTKKEPIAAG 770
 Qy 236 -IQD-GLAK-----GENDETTSNTL-----TLTNGLERRTK 264
 Db 771 KLKDKGVKVKIKKEGKTEAAATAGTAATAAAGVAAAGPAKLEAERSLMSSEPDLT 830
 Qy 265 TYSEDNFEEQLQYFNFYALLKSIDSEKAKE-----KETLIITIMKTLIDFVKMMV 314
 Db 831 DFELKAEIDVAKDIKPQLELIEDEKLEKTEPEGAUVIQTKEVS-----877
 Qy 315 KYGTISPGEVSVLENLDEMIALQTKNKL---EKQATDNISKL-----PPAPSEKSHHE 365
 Db 878 KGSAESDEGITTTEGESE--CEQTPPEEPVEKQGVDDIEKPEDEGAGFESSEAGDYE 935
 Qy 366 TDSKTEAAKWEKEYGSLKSDKDDNSNPGKTDPEKGTKEA--YLEAIRKNI-----416
 Db 936 EKAETEEAEEPEEDG-----EDNVSGSASKHSPTDEETIAKAEADVHIKKEKRESVAGSDORA 992
 Qy 417 -----EWLKKHDKKKNKEDYDLSKWRDFINKQADAYVEKGLDKKEAE 459
 Db 993 EEDMDLEALKEGAEQSEEGESEE-----DKAEDAREEDHEPDKTEAE 1035
 RESULT 8
 RA50 METJJA STANDARD; PRT; 1005 AA.
 ID RA50 METJJA
 AC Q58718;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR MJ1322.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococcaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RC MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Smith H.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U67572; AAB99331.1; -;
 DR PIR; A64465; A64465.
 DR TIGR; MJ1322; -;
 DR HAMAP; MF_00449; -; 1.
 DR InterPro; IPR003593; AAA ATPase.
 DR InterPro; IPR003439; ABC transporter.
 DR InterPro; IPR003405; SMC C.
 DR InterPro; IPR003395; SMC N.
 DR Pfam; PF04423; Rad50 zn hook; 1.
 DR Pfam; PF02483; SMC C_1.
 DR Pfam; PF02463; SMC N_1.
 DR ProDom; PD000006; ABC transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR DNA repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
 KW NP BIND 32 39
 FT ATP (BY SIMILARITY)
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 158 849
 SQ SEQUENCE 1005 AA; 119387 MW; 98BB48173E788F3 CRC64;
 Query Match 7.5%; Score 179; DB 1; Length 1005;
 Best Local Similarity 22.3%; Pred. No. 0.12;
 Matches 116; Conservative 80; Mismatches 164; Indels 160; Gaps 27;
 QY 52 DKIKKTPPENKPGQSNYSFVNLNLKA-ITKEKIEKQRQIRSPDLNKLN-VEVDV 109
 DB 159 DEFECY---QKMGIEVKEKRLERIEGELNYKENYKE-----LNKMSQLEE-- 205
 QY 110 STNKRKLIDDYDTKSLGDKHFQDDPLGLHQLDTPLTAEDIVHKLAARIYENDRAVFD 169
 DB 206 --KNKLMEINDKLNK-IKKEFEDIEKLFNENWKLLEYKFNKL-----BERKALEL 257
 QY 170 KIVS-KLLNLGLTESQAH-----TLEDAVAVLQKLIKSEANNYEEDPNK 214
 DB 258 KNGELKILEYDLNTVVEARETLNRHKDEVEKYKSLVDEIRKIESRL--RELKSHVEDYKL 315
 QY 215 PTGWTENQAGKIEKVTPMIAIODGLAKGNDVNTLTITNGLERRTKTSYSEDNFEEL 274
 DB 316 LTQLEILKDI-EK-----LKEFINKSKYRDDIDNLTDLNKL-----KDEIERV 360
 QY 275 QYFNFVALLKSIDSKAEK-----ETLITIMKTLIDFVKMVKYGTISPE 322
 DB 361 ETIKDLLEELKLNLEIEKIEKICEKYEYKYLEEKAVEYNKLTLEYITL--- 417
 QY 323 EGVSYLENLDMALQTKNLEKQVTD---NISKLPAPS-----EKSHEETDSTK-- 370
 DB 418 -----LOEKKSIENKINDLETRINKLEETKNIDIESIENSLKEIEKKV 463
 QY 371 -----PEAAKWEKEYGLKSDSTK-----DDNSNPGK-----TDEPK-----GKT 405

Db 464 LENLQKEKIELNKKLGEINSEIKRLKILDELKEVEGKPLCKTKPTIDENKQELINQHT 523
 QY 406 E--AYLE-----AIRKNIEMKKH-----DKKGKEDYDL----- 433
 Db 524 QLNKKYTELEENKKIRETEKIEKLEKIDKEENLTKLTLYLEKQSQIELELKLKRY 583
 QY 434 SKMRDPINKQADAYVEKG-----ILD--KEEABAIKRIYS 466
 Db 584 KEQLDEINKKISNYVINGKPVDEILEDIKSQLNKPKNFYN 623
 RESULT 9
 REST_HUMAN
 ID REST_HUMAN STANDARD; PRT; 1427 AA.
 AC P30622;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Restin (Cytoplasmic linker protein-170 alpha-2) (CLIP-170) (Reed-
 DE Sternberg intermediate filament associated protein).
 GN RSN.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX TISSUE=Peripheral blood monocytes;
 RC MEDLINE=92289675; PubMed=1600942;
 RA Bilbe G., Delabie J., Brueggen J., Richener H., Asseberg F.A.M.,
 RA Cretetti N., Sorg C., Odink K., Tarsay L., Wiesendanger W.,
 RA de Wolf-Petersen C., Shipman R.;
 RT "Restin: a novel intermediate filament-associated protein highly
 RT expressed in the Reed-Sternberg cells of Hodgkin's disease.";
 RL EMBO J. 11:2103-2113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92405160; PubMed=1356075;
 RA Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
 RT "CLIP-170 links endocytic vesicles to microtubules.";
 RL Cell 70:887-900(1992).
 CC -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
 CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES
 CC -!- SURCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
 CC CYTOSKELETON.
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P30622-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P30622-2; Sequence=VSP_000765;
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
 CC OF HODGKIN'S DISEASE.
 CC -!- SIMILARITY: Contains 2 CAP-Gly domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; X64838; CAA4050.1; -;
 DR EMBL; M97501; AAA35693.1; -;
 DR PIR; S22695; S22695.
 DR Genew; HGNC:10461; RSN.
 DR MIM; 179838; -;
 DR GO; GO:0005768; C:endosome; TAS.
 DR GO; GO:0005882; C:intermediate filament; TAS.
 DR GO; GO:0015630; C:microtubule cytoskeleton; TAS.
 DR GO; GO:0008017; F:microtubule binding activity; TAS.

```

DR GO: 0006899; P: non-selective vesicle transport; TAS.
DR InterPro: IPR000938; CAP-Gly.
DR PFam: PF01302; CAP_GLY_2.
DR SMART: SM00343; Znf_C2HC; 1.
DR PROSITE: PS00845; CAP_GLY_1; 2.
DR PROSITE: PS0245; CAP_GLY_2; 2.
KW Cytoskeleton; Microtubules; Coiled coil; Repeat; Alternative splicing.
FT DOMAIN 78 120 CAP-GLY 1.
FT DOMAIN 143 204 SER-RICH.
FT DOMAIN 232 274 CAP-GLY 2.
FT DOMAIN 304 331 SER-RICH.
FT DOMAIN 350 1342 COILED COIL (POTENTIAL).
FT DOMAIN 1408 1421 CCHC-BOX.
FT VARSPIC 457 491 Missing (in isoform Short).
FT CONFLICT 1069 1069 /FTD=VSP_000765.
FT SEQUENCE 1427 AA; 160989 MW; 0A4F166DD94254E8 CRC64;

Query Match
Best Local Similarity 22.4%; DB 1; Length 1427;
Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;

QY 34 RELSAG-----RPLNEQIAEAEEDKIKKTPPENKPGQSNYSFVDNL-----NLKA 80
DB 769 RASSEGKSEKKRQQL-EAEKQIRHLEIEKKAESSKASSITRELQGRLEKLTNIQEN 827
QY 81 ITE-----KEKIEKQ-----SIRSPDL-NKL-----NVEDVDSTKN 113
DB 828 LSEVSQVKETLEKELQILKEFAEASBEAVSQSMQETVNLKQKEQFNWLSDDLKL 887
QY 114 RKLDDYSTSGLDHFKQDDPGLHLDGTPLEAEDIVHKIARIYEENDRAVDFKIVS 173
DB 888 RENLAD-----MEAFKEREDEEQL-----IKAKEKLEND-----IA 920
QY 174 KLLNLGLITESQAHTLEDEV-----AEVLQKLISKEANNYE-----EDPNKPTSWTEN 221
DB 921 EIMKSGNSQLTWNDELKRDVEELQKLTKANENASFLQKSIDETVVAESQ 980
QY 222 QAGKIPKVTMAAIQDGLAGENDETVSNLTITLNGLERTKYTSNDFEELQYFPNFY 281
DB 981 EAAKHKEB-----EKLEKRLSD-----LEKQWET-SHNQCELEK----- 1015
QY 282 ALLKSIDSEKAKETLITIMKTLIDFVKRMVKYGTISPPEGVSYLENLDMIALQTKN 341
DB 1016 ARYERATSETTKHEEILQNLQKTLITLDTEDKLG-----ARENSGLLQLEELKQADKA 1071
QY 342 KLEKNATNISKLPAPSEKSH--EETDSTKEAAKMEKYGSLKSTKDDNSNPGKTD 399
DB 1072 KAAQTAEDAMQIMEQMTKEKTETTLASUEDTKTNAKLQNLDTLKNNL-KNVEELNKS 1130
QY 400 EPKGKTEAYLEAIRKNIEWLKKH-----DKGNKEDYDLSKWRDPINKQADAYV 448
DB 1131 ELLTVENQKMEFEKFEITLQAAQKSQLQALQENVKLAELGSRDEVTSHQKLEE 1190
QY 449 EKGILDKEEAFAIKR 463
DB 1191 ERSVLNNQLEMKR 1205

RESULT 10
RA50_PYRFU STANDARD; PRT; 882 AA.
AC P58301;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DE 15-SEP-2003 (Rel. 42, Last annotation update)
DE DNA double-strand break repair rad50 ATPase.
GN RAD50 OR PF1167.
OS Pyrococcus furiosus.
OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
OC Pyrococcus.
OX NCBI_TaxID=2261;

```

```

RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
RP STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=20485553; PubMed=11029422;
RA Hopfner K.-P., Karcher A., Shin D., Fairley C., Tainer J.A.,
RA Carney J.P.;
RT "Mrell and Rad50 from Pyrococcus furiosus: cloning and biochemical
RT characterization reveal an evolutionarily conserved multiprotein
RT machine.";
RL J. Bacteriol. 182:6036-6041(2000).
RP SEQUENCE FROM N.A.
RP STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
RT "The complete sequence of the Pyrococcus furiosus genome.";
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
RP X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-149.
RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
RX MEDLINE=20348838; PubMed=1092749;
RA Hopfner K.-P., Karcher A., Shin D.S., Craig L., Arthur L.M.,
RA Carney J.P., Tainer J.A.;
RT "Structural biology of Rad50 ATPase: ATP-driven conformational
RT control in DNA double-strand break repair and the ABC-ATPase
RT superfamily.";
RL Cell 101:789-800(2000).
CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
CC rad50/mrell complex possesses single-strand endonuclease activity
CC and ATP-dependent double-strand-specific exonuclease activity.
CC Rad50 provides an ATP-dependent control of mrell by unwinding
CC and/or repositioning DNA ends into the mrell active site.
CC -1- SUBUNIT: Forms a complex with mrell.
CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AE010225; ALU81291.1;
CC PDB; 1F2T; 20-SEP-00.
CC PDB; 1F2U; 02-AUG-00.
CC PDB; 1I18; 30-MAY-01.
CC PDB; 1L8D; 28-AUG-02.
CC HAMAP; MF_00449; -.
CC InterPro; IPR003439; ABC transporter.
CC InterPro; IPR003395; SMC N.
CC Pfam; PF04423; Rad50 zn hook; 1.
CC Pfam; PF02463; SMC N; 1.
CC ProDom; PD000006; ABC transporter; 1.
KW DNA repair; Hydrolase; ATP-binding; Coiled coil; 3D-structure;
KW Complete proteome.
FT NP BIND 30 37 ATP.
FT DOMAIN 148 744 COILED COIL (POTENTIAL).
SQ SEQUENCE 882 AA; 103839 MW; 3ADCBD250382A99E CRC64;

Query Match
Best Local Similarity 22.1%; DB 1; Length 882;
Matches 112; Conservative 90; Mismatches 172; Indels 132; Gaps 21;

QY 39 ERLNEQIAEAEER-DKIK---KTPPENKPGQSNYSFVDNLNLKA-----ITEKIE 88
DB 260 KRTLEERIKNTBELYELKLEKEKEBELBEQVKETISIKKQVDAYLALKEFKNEYLDKYYKE 319
QY 89 KERQSISSPLDNKLVNVEDVDSTKNKRLIDDYSTKSGLDHFKQDDPGLHLDGTPLEA 148
DB 320 KE--LTFVBEELINEIQ-----KRIEELNEKESEKEKLENE-----K 353
QY 149 EDIVHKIARIYEENDRAVDFKIVSKLLNLGLITESQAHTLEDEVAEVLQKLISKEANNY 208
DB 149 EDIVHKIARIYEENDRAVDFKIVSKLLNLGLITESQAHTLEDEVAEVLQKLISKEANNY 208

```

Db 354 KEILNKLAIA---EKDQHYEEIKAKENLRQLKEKLGDKSPEDIKLLEETK-ITI 409

QY 209 BEDPNKPTSWTENQAKIPEKTPMAAQDGL--AKGE-----NDE----- 247

Db 410 EERNEIT---ORIGELKNKIGDGLKTAIEELKKAAGKPCVCGRELTDREELLSKYLH 465

QY 248 -----TVSNLTITLNGLERTKTYSNDNFPELQVFNFYALLKSID- 290

Db 466 DLNNSKNTLAKLIDRKSLELERLIDME--IKLTPLTVAEQIRSIIEELNVNLEKIE 524

QY 291 KEAKEKETLITIMTKLIDFVKQKWKYGTISPEBGVSYLENIDEMIA--LOTNKLKLNAT 348

Db 525 KNATYELKLEELRLEGRIGLA-----EDLKLAPLEKLAALIHKKQSELEKELK 576

QY 349 DNISKLFPAPSKSHEETDSTKEAAKWEKVGSLKDKTD- 399

Db 577 ELNTKL--ESFGKSVEDLSDKLELEIYKRYLTLLNSKKELEITQREIAKAKETLEMSF 635

QY 400 EPKGTAYLEAIRKNIWELKHKDKKNKEDY-----DLKMR 437

Db 636 BELAEVADIERIEKLSQLK--QKYNEEYKXKREKELEKELARLEAKQKELEKRR 692

QY 438 DFINKQADAYVEKGLDKKEEAIAKR 463

Db 693 DTIK-----STLEKLAKEKNERVKK 714

RESULT 11

BAG_STRAG STANDARD; PRT; 1164 AA.

AC 01-AUG-1992 (Rel. 23, Created)

DT 01-AUG-1992 (Rel. 23, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE IGA FC receptor precursor (Beta antigen) (B antigen).

GN BAG

OS Streptococcus agalactiae.

OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;

OC Streptococcus.

OX NCBI_TaxID=1311;

RN [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-48.

RC STRAIN=LA239;

RX MEDLINE=91312121; PubMed=1857207;

RA Jeristrom P.G., Chhatwal G.S., Timmis K.N.;

RT "The Iga-binding beta antigen of the c protein complex of Group B

RT streptococci: sequence determination of its gene and detection of two

RT binding regions";

RL Mol. Microbiol. 5:843-849(1991).

RN [2]

RP IDENTIFICATION OF IG-LIKE DOMAIN.

RX MEDLINE=97035265; PubMed=8880921;

RA Bateman A., Eddy S.R., Chothia C.;

RT "Members of the immunoglobulin superfamily in bacteria.";

RL Protein Sci. 5:1939-1942(1996).

CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by

CC an amide bond (potential).

CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to license@isb-sib.ch).

CC -----

CC EMBL; X59771; CAA42442.1; ..

DR PIR; S15330; FCSOAG. Csurface antigen.

DR InterPro; IPR004829; Csurface antigen.

DR InterPro; IPR005877; Gpos YSIRK.

DR InterPro; IPR001899; Gram_pos_anchor.

DR InterPro; IPR003599; Ig.

InterPro; IPR006192; LPXTG.

Pfam; PF00746; Gram_pos_anchor; 1.

Pfam; PF05062; RICH; 1.

Pfam; PF04650; YSIRK signal; 1.

ProDom; PD153432; Csurface_antigen; 1.

SMART; SM00409; IG; 1.

TIGRfams; TIGR01167; LPXTG anchor; 1.

TIGRfams; TIGR01168; YSIRK signal; 1.

ProSITE; PS00847; GRAM_POS_ANCHORING; 1.

ProSITE; PS00835; IG LIKE; FALSE NEG.

Cell wall; Peptidoglycan-anchor; Receptor; Repeat; Signal;

Immunoglobulin domain.

FT SIGNAL 1 37

FT CHAIN 38 1135 IGA FC RECEPTOR.

FT PROPEP 1136 1164 REMOVED BY SORTASE (POTENTIAL).

FT DOMAIN 434 534 IG-LIKE.

FT DOMAIN 199 438 IGA-BINDING (POTENTIAL).

FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).

FT DOMAIN 827 945 PRO-RICH REPEATS.

FT SITE 1132 1136 LPXTG SORTING SIGNAL (POTENTIAL).

FT MOD_RES 1135 1135 AMIDE-LINKED TO CELL WALL (POTENTIAL).

FT SEQUENCE 1164 AA; 131051 MW; 65DE94AF720A5474 CRC64;

Query Match 7.2%; Score 172.5; DB 1; Length 1164;

Best Local Similarity 21.6%; Pred. No. 0.3; 189; Indels 185; Gaps 25;

Matches 124; Conservative 77; Mismatches 77

QY 36 LSAERPLNE-----QIAEAEDKIKITPPENKPGQSNYSFVDNLNLKAIT-EKEK 86

Db 149 LLENQFNETNRLHLIKQHEEVEKDKKAK---QOKTLKQSDTKVDLSNIDKELNHQSQ 204

QY 87 IEK--ERQSISSPLDNKI-----NVEDVDSTKNRKLIDDDYSTKGLDHLK 130

Db 205 VEKMAEQGITNEDKDSMLKKIEDIRKQAQADKEDAEVKVREBELGLFSSTKAGLDQE 264

QY 131 FQDDPDGLHOLDGTPLTAEIDVHKI-----AARIYEENDRAVDFDKIVSKLLNLG 179

Db 265 IQE-----HVKKET--SSEENTQKDEHYANSLQNLAKSLEELDATTNQATQVKNQF 317

QY 180 L-----ITESQAHLEDEVAEV-----LQKLISKE---ANNYEE 210

Db 318 LENAQKKEIQPLIKETNVKLYKAMSESLQVEKELKHNSANLEDLVAKSEIVREYEG 377

QY 211 DNKETSWE-----NQAGKIPEKVTMAAIODGLAKGENDETVSN 251

Db 378 KLNQSKNLPKQLEEEAHSLKQVDFRKKFTSQVTPKRVYGRDLAANENNQ---Q 434

QY 252 TLTLNGLERTKTYSEDNFEELQVFPNFYALLKSIDSEKAEKETLITIMTKLIDFVK 311

Db 435 KIELTVSPENITVYEGED-----VKFTVTAKS-DS-----KTLDFSD 471

QY 312 MMVKYG-TISPEBGVSYLENLD-----EMIALQTK-----NKLEKNATD 349

Db 472 LLTKYNPSVSDRISTNYKNTNDNHKIAEITIKNLKLNESQTVTLKAKDDSGNVVEKFTI 531

QY 350 NISKLFPAPSKSHEETDSTKEAA-----KMEKEYGSLKD 385

Db 532 TVQKKEKQVPTQKDSKTEEKVQPEPSNDKQQLQELIKSAQOELEKLEKAKELME 591

QY 386 STKDDNSNPGKTDPEKGTAYLEAIRKNIEMLKK-----HDKKGNKEDYDLSKMRFIN 441

Db 592 Q-PEIPSNP--EYGIQKSIWESQKEPIQEAITSFKIIGDSSSSKYTYEHYFNKYKSDFMN 648

QY 442 KQADAYVEKGLDKKEAE-----AIKRIYSS 467

Db 649 YQLHAQME--MLTRKVVQVMNKPVDNAEIKKIFES 681

RESULT 12

TRDN CANFA

ID TRDN CANFA STANDARD; PRT; 700 AA.

AC P82179;

DT 28-FEB-2003 (Rel. 41, Created)

Qy	83	EKEKIEKRSIRSPDLNKLNVEDVSTQNRKLLDDYDVSYSGLDHHFQDDPDDLHQLD	143
Dq	:	:	:
Dd	182	HKEKIEKKRP-----ETKTMAKEERKAkteEKIK--KEVKGKGQEKVKPTAAKVKEVQ	233
Qy	143	GTPLTARDIVHKIAARI--YEENDRAVFDDKIIVSKLLNLGLITESQAHTL-	189
Dd	:	:	:
Dd	234	KTPPKAKEKEGKETAAVAKHQKOQYAFCRYMIDMFVHGDLRPGQSPALPPPLPTVQASR	293
Qy	190	-----EDEVAEVLQKLISKEANNYYBEDPNKPPTS-----WTENQAGKIPE---	228
Dd	:	:	:
Dd	294	PTPASPTLEGKEBEBEKKAEKKVTSETKKEKEDVKKSDDKDTAIDVEKKEFGKAPETKQ	353
Qy	229	---KVTPMALIQDLAKGENDETVSNITLTNLNGELRRRTKYTSYEDNFEELOYPNPVALLK	285
Dd	:	:	:
Dd	354	GTIKVVAQAAA-K--KDEKEDSKTKT- --PVEEHPKGGKQEKKE-----K	394
Qy	286	SIDSEKAEKAKETLITIMKTLIDFKVMVKYGTISPREGSVYLENDLMIALQTNNKLEK	345
Dd	:	:	:
Dd	395	YVEPAKSSKKEHS-----APSE-----KQVKAKTERAKEE	424
Qy	346	NATDNISKLFPAPSEKSHEETDSTKEEAAMKEGYSLKSDTKDONSPPGGKTDE--PKG	403
Dd	:	:	:
Dd	425	TSAASTKAVGP---KKEETKTIVEQEIRKEK---SGKTSTASKDKKEPEIKKDEMPKA	478
Qy	404	KTEAYLEAIRKNIEWLKKHD---KKGNKEDVDLSKMRDFIN-----KQADAYVEKG	451
Dd	:	:	:
Dd	479	DKEVKPPQSQVKKESSESQVKEAKPEODIAPKPTVSHGKPEEKVWKQVKA--TEKA	537
Qy	452	ILDK 455	
Dd	:		
Dd	538	AIEK 541	
RESULT 13			
KF5C_HUMAN			
ID	KF5C_HUMAN	STANDARD;	PRT; 957 AA.
AC	O60282; O95079;		
DT	16-OCT-2001 (Rel. 40, Created)		
DT	16-OCT-2001 (Rel. 40, Last sequence update)		
DT	16-OCT-2001 (Rel. 40, Last annotation update)		
DE	Kinesin heavy chain isoform 5C (Kinesin heavy chain neuron-specific		
DE	2).		
GN	KIF5C OR NKHC2 OR KIAA0531.		
OS	Homo sapiens (Human).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Primates; Carnivora; Hominidae; Homo.		
OX	NCBI_TaxId=9606;		
RN	[1]		
RN	SEQUENCE FROM N.A.		
RC	TISSUE=Brain;		
RC	MDLINE=98290545; PubMed=9628581;		
RA	Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,		
RA	Nomura N., Ohara C.;		
RT	"Prediction of the coding sequences of unidentified human genes. IX.		
RT	The complete sequences of 100 new cDNA clones from brain which can		
RL	code for large proteins in vitro.";		
RL	DNA Res. 5:31-39(1998).		
RN	[2]		
RP	SEQUENCE OF 355-585 FROM N.A.		
RA	Engelender S., Sharp A.H., Colomer V., Tokito M.K., Lanahan A.,		
RA	Worley P., Holzbaur E.L.F., Ross C.A.;		
RT	"Huntingtin associated protein 1 (HAP1) interacts with the p150Gluc		
RT	subunit of dynactin.";		
RL	Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.		
CC	-1- FUNCTION: KINESIN IS A MICROFILAMENT-ASSOCIATED FORCE-PRODUCING		
CC	PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.		
CC	-1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT		
CC	CHAINS.		
CC	-1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN, PROSTATE, AND		
CC	TESTES, AND MODERATE EXPRESSION IN KIDNEY, SMALL INTESTINE, AND		
CC	OVARY.		
CC	-1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-		
CC	TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF		

CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
 CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
 CC DIMERIZATION; AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
 CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
 CC VESICLES AND MEMBRANOUS ORGANELLES.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC EMBL; AB011103; BAA25457.1; -;
 CC EMBL; AF010146; AAD01436.1; -;
 CC HSP; P56536; 2KIN.
 CC Genew; HGNC:6325; KIF5C.
 CC MIM; 604593; -;
 CC GO; GO:0005871; C:kinesin complex; TAS.
 CC GO; GO:0003777; F:microtubule motor activity; TAS.
 CC GO; GO:0006996; P:organelle organization and biogenesis; TAS.
 CC InterPro; IPR001752; kinesin_motor.
 CC Pfam; PF00225; kinesin; 1.
 CC PRINTS; PRO0380; KINESINHEAVY.
 CC SMART; SM00129; KISG; 1.
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 CC PROSITE; PS00667; KINESIN MOTOR DOMAIN2; 1.
 CC Motor protein; microtubules; ATP-binding; Coiled coil.
 CC DOMAIN 1 386 KINESIN-MOTOR (BY SIMILARITY).
 CC FT DOMAIN 406 923 COILED COIL.
 CC FT DOMAIN 859 956 GLOBULAR.
 CC FT DOMAIN 174 315 MICROTUBULE-BINDING.
 CC FT NP BIND 86 93 ATP (BY SIMILARITY).
 CC FT CONFLICT 355 360 TLKNI -> STHASV (IN REF. 2).
 CC FT CONFLICT 583 585 EFT -> DRV (IN REF. 2).
 CC SEQUENCE 957 AA; 109494 MW; A9F25B81C994322A CRC64;
 CC
 CC Query Match 7.0%; Score 167.5; DB 1; Length 957;
 CC Best Local Similarity 20.3%; Pred. No. 0.42;
 CC Matches 103; Conservative 98; Mismatches 182; Indels 125; Gaps 23;
 CC
 CC QY 25 GSQKSLNR-----ELSAERPLNEQIAEAEDEKIKKTPPENKPGOSYFVDNLNLL- 78
 CC DB 321 GQRAKTIKTVSNVLELTAE-----EWKKYKEKEKNKTLKNVIQHLMEEL 367
 CC QY 79 -----KAI TEKEKIEKERSIRSSPLDNKLVEDVD-----STKNRKLIDYDSTKGG 126
 CC DB 368 NRWNGEAVPEDEQISAKDQK-NLEPCDNPIDNIAPVVAGISTBEK---EKYDEEIS 423
 CC QY 127 LDHKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEEND-----RAVFDKIVSKLLNLGLI 181
 CC DB 424 LYRQLDKDDEINQ-----QSQAELKQQLMDQDDELLASTRDYKEIKQELTFLQTE 476
 CC QY 182 TESAHTLEDEVAELVKLSKEANNYE-----EDNKPTSTWENQAKIKPEKVPMAAIQD 238
 CC DB 477 NEA-----AKDEVKVLQAL-EELAVNYDQKSQEVEDKTRANEQLTDELAQKTTTLTTQ 531
 CC QY 239 GLAK-----GENDETVSNTLTTLTNGLEER-----TK 264
 CC DB 532 ELSQQLSNHKKRATETIINLLKDLGEGGIIGND---VKTLADVNGVIEEFTMAR 588
 CC QY 265 TYSEDNFEELQYFPNFALLKS---IDSEKAEKETLITIMKTLDIFVQVMKYGTISPE 322
 CC DB 589 LYISKMKSEKSVLNRSKQLESQMSQMSNRKWNASERELAAQQLLISQHEAKISLT---- 644
 CC QY 323 EGVSYLENIDEMALQTKNLEK---NATDNTSKLPAPSEKSHETDSTKE----- 371
 CC DB 645 ---DYQNMW-----QKRROLEESQDSLSLEAKL---RAQERKMHEVSFQDKEHLRLIQ 694
 CC QY 372 EAAKMEKEYGSLKDDSTKDDSNPNPGGKTDEPKGKTEAYLEAIRKNTEWLKHKDKGKNEDY 431

DB 695 DAEMKKALEQQWESHREAHQKLSRLDRDEIEBKQKIIDEIRLNKQJLEQEK-LSSDY 753
 QY 432 DLSKMRDFINKQADAYVEKGL--DKEE 457
 DB 754 NKLKIED---QEREMKLEKLLNDKRE 778
 RESULT 14
 MLPI_YEAST
 ID MLPI_YEAST STANDARD; PRT; 1875 AA.
 AC Q02455;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin-like protein MLPI.
 GN MLPI OR YKR095W OR YKR415.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=93247549; PubMed=8483450;
 RA Koelling R., Nguyen T., Chen E.Y., Botstein D.;
 RT "A new yeast gene with a myosin-like heptad repeat structure.";
 RL Mol. Gen. Genet. 237:359-369 (1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=94205265; PubMed=8154186;
 RA Rou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
 RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
 RT "The complete sequence of a 15,820 bp segment of Saccharomyces
 RT cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three
 RT new open reading frames.";
 RL Yeast 9:1349-1354 (1993).
 CC -!- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
 CC REPAIR.
 CC -!- SIMILARITY: SOME, TO THE TPR ONCOGENE.
 CC -!- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL1".
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch)
 CC
 CC EMBL; L01992; AAA34783.1; -;
 CC EMBL; X73541; CAA51948.1; -;
 CC EMBL; Z28320; CAA82174.1; -;
 CC PIR; S38173; S38173.
 CC SGD; S0001803; MLPI.
 CC DR GO; GO:0005635; C:nuclear membrane; IDA.
 CC DR GO; GO:0005654; C:nucleoplasm; IDA.
 CC DR GO; GO:0006606; P:protein-nucleus import; IDA.
 CC KW Coiled coil; DNA repair.
 CC FT DOMAIN 69 487 COILED COIL (POTENTIAL).
 CC FT DOMAIN 531 1678 COILED COIL (POTENTIAL).
 CC FT DOMAIN 1834 1866 COILED COIL (POTENTIAL).
 CC FT CONFLICT 301 301 R -> A (IN REF. 1).
 CC SEQUENCE 1875 AA; 218455 MW; 683A0D34C9066867 CRC64;
 CC
 CC Query Match 6.9%; Score 166.5; DB 1; Length 1875;
 CC Best Local Similarity 23.6%; Pred. No. 1;
 CC Matches 112; Conservative 85; Mismatches 178; Indels 99; Gaps 28;
 CC
 CC QY 30 SLHNRSLAERPLNEQIAEAEDEKIKKTPPENKPGOSYFVDNLNLLKAITE-KEKIE 88
 CC DB 588 TLKSEKMDLESRIEELQKELEE--LKTSVPNE-----ASISNV-----TIKQLETETKRDLE 637

Qy 462 KRIYSL 468
Db 1596 KRVQAL 1602

Search completed: January 2, 2004, 12:34:53
Job time : 19 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 12:33:21 ; Search time 40 Seconds
(without alignments)
3019.215 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGTTWLVLVLPQAF.....EKGILDKKEAEAKRIYSSL 468

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SPTREMBL_23.*

- 1: sp_archaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_organelle.*
- 9: sp_phase.*
- 10: sp_plant.*
- 11: sp_rodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_virus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2062	86.0	466	11 Q8R1D7	Q8R1D7 mus musculus
2	1216	50.7	457	13 Q91847	Q91847 xenopus lae
3	198	8.3	7210	5 Q9V7G8	Q9V7G8 drosophila
4	198	8.3	9270	5 Q8MLD9	Q8MLD9 drosophila
5	197	8.2	2081	10 Q9LH98	Q9LH98 arabidopsis
6	196.5	8.2	1175	16 Q8XNW6	Q8XNW6 clostridium
7	196	8.2	2139	5 Q07569	Q07569 entamoeba h
8	191.5	8.0	2760	5 Q815Y2	Q815Y2 plasmodium
9	189	7.9	1434	5 Q81492	Q81492 plasmodium
10	188.5	7.9	1661	5 Q06166	Q06166 plasmodium
11	184	7.7	951	5 Q96229	Q96229 plasmodium
12	183.5	7.6	1790	3 Q07380	Q07380 saccharomyc
13	183.5	7.6	2612	5 Q815X5	Q815X5 plasmodium
14	180.5	7.5	3504	5 Q81L45	Q81L45 plasmodium
15	179.5	7.5	1320	11 Q9JK25	Q9JK25 rattus norv
16	179.5	7.5	1510	5 Q25920	Q25920 plasmodium

17	179	7.5	853	10 Q9LFE4	Q9LFE4 arabidopsis
18	179	7.5	1558	5 Q96275	Q96275 plasmodium
19	179	7.5	1786	5 Q9U0P0	Q9U0P0 plasmodium
20	179	7.5	2269	5 Q26223	Q26223 plasmodium
21	179	7.5	2747	5 Q9BJX9	Q9BJX9 plasmodium
22	178.5	7.4	1323	5 Q9NB35	Q9NB35 plasmodium
23	178.5	7.4	3317	16 Q8EWP8	Q8EWP8 mycoplasma
24	178	7.4	840	5 Q815X4	Q815X4 plasmodium
25	177.5	7.4	1464	5 Q81IF6	Q81IF6 plasmodium
26	177	7.4	1112	16 Q8EWQ1	Q8EWQ1 mycoplasma
27	177	7.4	1871	10 Q9SRD5	Q9SRD5 arabidopsis
28	176.5	7.4	682	5 Q81BV6	Q81BV6 plasmodium
29	175.5	7.3	1964	5 Q8SWQ7	Q8SWQ7 loligo peal
30	175	7.3	10578	5 Q81SF5	Q81SF5 caenorhabdi
31	175	7.3	18519	5 Q81SF6	Q81SF6 caenorhabdi
32	175	7.3	18534	5 Q81SF7	Q81SF7 caenorhabdi
33	174.5	7.3	924	5 Q15738	Q15738 dictyosteli
34	174.5	7.3	1263	5 Q81K49	Q81K49 plasmodium
35	173.5	7.2	1795	5 Q81E35	Q81E35 plasmodium
36	172.5	7.2	1134	2 Q99051	Q99051 streptococc
37	172.5	7.2	1979	5 Q96133	Q96133 plasmodium
38	172	7.2	495	5 Q81EK6	Q81EK6 plasmodium
39	171.5	7.1	2083	5 Q9N435	Q9N435 caenorhabdi
40	171.5	7.1	3484	5 P91257	P91257 caenorhabdi
41	171.5	7.1	5507	5 Q81HN3	Q81HN3 plasmodium
42	170.5	7.1	2771	5 Q26216	Q26216 plasmodium
43	170	7.1	1140	4 Q9ULE5	Q9ULE5 homo sapien
44	170	7.1	1804	11 Q921Q1	Q921Q1 mesocricetu
45	170	7.1	3130	5 Q9BK46	Q9BK46 plasmodium

ALIGNMENTS

RESULT 1

Q8R1D7 ID Q8R1D7 PRELIMINARY; PRT; 466 AA.
AC Q8R1D7;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Similar to secretogranin III.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC024785; AAH24785.1; -
SQ SEQUENCE 466 AA; 52783 MW; AB1CB9667276315C CRC64;

Query Match	86.0%;	Score 2062;	DB 11;	Length 466;
Best Local Similarity	87.0%;	Pred. No. 2.1e-94;		
Matches 410;	Conservative 17;	Mismatches 36;	Indels 8;	Gaps 2;
Qy	1	MGFLGTGTTWLVLVLVLV---	PIQAPFKPGSQSKSLHNREL	SAERPLNEQIAPAEADKIKKT 57
Db	1	MGFLWTGTSWLVLVLSGPIQAF	PKPGSQSKSLHNREL	SAERPLNEQIAPAEADKIKKA 60
Qy	58	YPPEKPKGQSNYSVDNLNLLKA	ITEKEKTEKERSIRSSPLDNK	LAVEDVDSVTKNRKLI 117
Db	61	FPSEKFSSESYSSVDNLNLLRA	ITEKETVEKEKQSIIRSPFD	NQNLNVEDADSTKNRKL 120
Qy	118	DDYDSTKSGLDHKFQDDPDGL	HLQJLDGTLTAEDIVHKIAA	RIYBENDRAVDFKIVSKLLN 177
Db	121	DEYDSTKSGLDHKFQDDPDGL	HLQJLDGTLTAEDIVHKIAA	RIYBENDRAVDFKIVSKLLN 180
Qy	178	LGLITESQAHTLEDEVAEVLQ	KLISKANNYEDPNKPTSWTEN	QAQKIPKVTPTMAAIQ 237
Db	181	LGLITESQAHTLEDEVAEALQ	KLISKANNYETLDKPTSTEN	QDQDKGIPEKVTPTVAAVQ 240

```

QY 238 DGLAKGENDTVSNTLTNGLRRTKTSYSDNPEELQYFNFYALLKSIDSEKEAKE 297
DB 241 DQFTNRENDTVSNTLTNSGLRRTNHRDEDFEELQYFNFYALLTSIDSEKEAKE 300
QY 298 TLITIMKTLDIFVKQWKYGTISPPEGVSYLENDEMIALQTNKLEKNATDNISKLFFA 357
DB 301 TLITIMKTLDIFVKQWKYGTISPPEGVSYLENDEMIALQTNKLEKNATDNISKLFFA 357
QY 358 PSEKSHETDSTKEAAKWEKEYGSLKDTKODNSNPGCKTDEPKGTEAYLEAIRKNIE 417
DB 356 PPEKSQETDSTKEAAKWEKEYGSLKDTKODNSNPGCKTDEPKGTEAYLEAIRKNIE 415
QY 418 WLKHKHKKGNKEDYDLSKMRDFINQADAYVEKGILDKKEAEAIKRIYSSL 468
DB 416 WLKHKHKKGNKEDYDLSKMRDFINQADAYVEKGILDKKEAEAIKRIYSSL 466

RESULT 2
Q91847
ID Q91847 PRELIMINARY; PRT; 457 AA.
AC Q91847;
DT 01-NOV-1996 (TrEMBLrel. 01, Created)
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE Secretogranin III.
OS Xenopus laevis (African clawed frog).
OC Amphibia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Int. pituitary;
RX MEDLINE=96217355; PubMed=8632145;
RA Holthuis J.C.M.; Martens G.J.M.;
RT "the neuroendocrine proteins secretogranin II and III are regionally
RT conserved and coordinately expressed with proopimelanocortin in
RT xenopus intermediate pituitary."
RL J. Neurochem. 66:2248-2256(1996).
DR EMBL: X92872; CAA63478.1;
SQ SEQUENCE 457 AA; 51903 MW; B3095E12D4B076D7 CRC64;

Query Match
Best Local Similarity 50.7%; Score 1216; DB 13; Length 457;
Matches 269; Conservative 63; Mismatches 106; Indels 34; Gaps 10;

QY 8 TWILVLPIQAPKPGSGODKSLHNRRLSARPLNEQIAEEDKIKTYPPENKPGQS 67
DB 9 TAIGFVHVQGFNPSIKQDKDVHNRLESEERPLEEQIAEA--DTVSRGTSTENQEMLK 66
QY 68 NYSFVDMNLKKAITEKEKIEKQSRSSPLDNKLVNVEDVSTKNRKLIDDDYDSTKSL 127
DB 67 NDSFADDLGLSMNGEKGDVKSSTGSGFSGNRGTLDADSTKRLKLAEPDITKSWP 126
QY 128 DHKQDDPDGLHOLDGTFPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGLITESQAH 187
DB 127 DYKYEDPDGLHOLD-APFTAEDIVRKIATRIYEENDRGVDFKIVSKLNLGLITESQAY 185
QY 188 TLDEVAEVLQKLSKEANNYEE-----DPNKTSWTENQAGIKPKVTFMMAIQDG 239
DB 186 TLDEVAEVLQQLIAEAKNREAGDLDSVRSVDYDDDKREGQW-----ET 234
QY 240 LAKGENDT--VNTLTNGLRRTKTSYSDNPEELQYFNFYALLKSIDSEKEAKE 297
DB 235 LNKNEDESSETKNEDSLST--VERNELSPED--DLQYFNFNRVLKSLNSENDVKEKK 290
QY 298 TLITIMKTLDIFVKQWKYGTISPPEGVSYLENDEMIALQTNKLEKNATDNISKLFFA 357
DB 291 TLITIMKTLDIFVKQWKYGTISPPEGVSYLENDEMIALQTNKLEKNATDNISKLFFA 347
QY 358 PSEKSHETDSTKEAAKWEKEYGSLKDTKODNSNPGCKTDEPKGTEAYLEAIRKNIE 416
DB 356 PPEKSQETDSTKEAAKWEKEYGSLKDTKODNSNPGCKTDEPKGTEAYLEAIRKNIE 415

```

```

DB 348 PSEKSHETDSTKEAAKWEKEYGSLKDTKODNSNPGCKTDEPKGTEAYLEAIRKNIE 405
QY 417 EWLKHKHKKGNKEDYDLSKMRDFINQADAYVEKGILDKKEAEAIKRIYSSL 468
DB 406 EWLKHKHKKGNKEDYDLSKMRDFINQADAYVEKGILDKKEAEAIKRIYSSL 457

RESULT 3
Q9V7G8
ID Q9V7G8 PRELIMINARY; PRT; 7210 AA.
AC Q9V7G8;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE CG18255 protein.
GN STRN-MLCK OR CG8304 OR CG18255.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7221;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkelley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbavani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
RA Burkis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Svangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fostel C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Hernandez J.R., Houck J.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banson J., An H., Baldwin D., Banson J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorset V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frise E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez C., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,

```



```
Db 1009 --NREKKEYEKSKTBEAKKESQDKKREKDSERKSKKEKESRDLKAKKKEE 1066
Qy 401 PKGKTEAYLEAIRKNIEWLKHDKGNKEDYDLKQKRDFFINKQADYVVEKGLDKEE 457
Db 1067 TKEKES-----ENHKKKKKKKHEHDNKSMMKEED--KKEKKKHEKSRKKEE 1115

RESULT 6
ID Q8XNW6 PRELIMINARY; PRT; 1175 AA.
AC Q8XNW6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2003 (TREMBlrel. 23, Last annotation update)
DE Probable exonuclease.
GN SBCC OR CPE0216.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_TaxID=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
EX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Hirakawa K., Yamashita A.,
RA Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater."
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001 (2002).
DR EMBL; AP003185; BAB79922.1; -.
DR InterPro; IPR003439; ABC_transporter.
KW Exonuclease; Complete proteome.
SQ SEQUENCE 1175 AA; 136878 MW; 7C6D2366525019C1 CRC64;

Query Match 8.2%; Score 196.5; DB 16; Length 1175;
Best Local Similarity 22.6%; Pred. No. 0.078;
Matches 116; Conservative 83; Mismatches 196; Indels 119; Gaps 21;

Qy 29 KSLHNRELSAERPLNEQIAEAEDEKIKTYPENKPGOSNYSFVDNL-NLLKAITEKE-- 85
Db 256 KELYDKRIEE-----SLVSRSEIKSFERVEISNKADKIVFINNLEILKEINKEDLK 311
Qy 86 --KIEKQSTRSSPLDNKLVNEDVDSTGNKRLIDDYSTKSGLDHKPDGPDGLHQLD 143
Db 312 FSELNKLLEELNLEENKLFEBFTKKKEKLPDLRLKKEKLLS--QKRDILFOIKA 369
Qy 144 TPLTAEDIVHKI-----AARIYENDRAVPDKIVSK-----LLNLG 179
Db 370 DGVKLKEACKKIFEDRSKCDTKLNSIEENKRLNEELKEEKEERKEELFVHEEFKNKINS 429
Qy 180 LITESQAHTLEDEVAEVLQKLI--SKEANNYEEDPNKPTSWTNOAGKIPKVTMAAIQ 237
Db 430 LFILNSYSLQKQNEIKSEVELKYYIKNTDEKSEK-----DLKVVESLSKIR 482
Qy 238 DGLAK-----GENDETVSNTLTTLNGLERTKTYSEDNFEELQYF 277
Db 483 DKLESLLKETPGDSNSILEKQIKLGEYREKLNKYKEIKNSLEESLKT--KNPFSEKLTFF 540
Qy 278 PNFYALLKSIDSEKAEKETLITIMKT--LIDFVMMVKYGTISPREGVSYLE----- 329
Db 541 ENQKLL-----EKEVRELKDYINKVVELAHLKRLNLEGECCPVCVSGTHHELNVKVE 595
Qy 330 -NLDE---MIALQTKNLEKNATDNISKL--FPAPSEK-----SHEETDSTKEBAK 375
Db 596 INLESNEKTIILSKSEKELILEFSKIEATLEYENKKEELNLEISIEEVEGVNEERLK 655
Qy 376 -MEKEYGSLKDTKDDNSNPGKTDPEPKGTAYLEAIRKNIEWLKHDKGNKEDYDL 434
Db 656 FLEEFNTLKDIEFNLK-----ENLEKDLKXK--EKKNNLE----- 693
Qy 435 KMRDFINKQADYVVEKGLDKEEAEAKIYSS 468
Db 694 ---NIFNKAIEVILCEKIVREKEIASKIKELDKEL 724
```

RESULT 7

```
Qy 007569 PRELIMINARY; PRT; 2139 AA.
AC 007569; 002504;
DT 01-NOV-1996 (TREMBlrel. 01, Created)
DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)
DT 01-OCT-2002 (TREMBlrel. 22, Last annotation update)
DE Myosin heavy chain.
GN MHCA.
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_TaxID=5759;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HMI:IMSS;
EX MEDLINE=93295430; PubMed=8515774;
RA Raymond-Denise A., Sansonetti P., Guillen N.;
RT "Identification and characterization of a myosin heavy chain gene
RT (mhca) from the human parasitic pathogen Entamoeba histolytica."
RL Mol. Biochem. Parasitol. 59:123-131 (1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HMI:IMSS;
RA Guillen N.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; L03534; AAB48065.1; -.
DR HSP; P08799; INMD
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00242; MYSC; 1.
DR SHARIT; 2139 AA; 245225 MW; C68307341DB51DD1 CRC64;
SQ SEQUENCE 2139 AA; 8.2%; Score 196; DB 5; Length 2139;
Best Local Similarity 21.0%; Pred. No. 0.17;
Matches 107; Conservative 108; Mismatches 191; Indels 104; Gaps 23;

Qy 20 FPKPGSQDKSLHNRELSAERPLNEQIAEAEDEKIKTYPEN-----KPGOSNYSFV 72
Db 859 FEEBEGKKKDEIED-----LKKKLAEE--EIKKREAAENALASATAKTGLEAKIQ 906
Qy 73 D-----NLLKAITEKEKIEKERSIRSSPLDNKLVNEDVD- 109
Db 907 DLEDKTISELSKLSAAELDKQELNLTENLEEDKEELKTDNLKGLDKSLKGLDLEV 966
Qy 110 -----STKNRKLIIDDYSTKSGLDHKPDGPDGLHQLDGLPTAEDIVHKAARIY 160
Db 967 ETELNSQIINTLN-ATVDKDKTIAEQESIDEKEDEITKLGDIKLEEEKDDL----- 1020
Qy 161 EENDR-----AVFPKIVSKLLNLGLITESQAHTLEDEVAEVLQKLI-SKEANNYEEDPNKPT 216
Db 1021 -EQDRADVSA TKDDIAKKLNTI-----ECEDAKDEIAKLEQLEDEE--NKNKDLTNEL 1073
Qy 217 SWTENQAKTIPEKVTMAAI-----ODGLAG--ENDETVSNTLTTLNGL-ERRTKTYS 267
Db 1074 QOTQLKGETEKSLAAQVAATKASDERDTLSQNLNLEKLTITKNLTKTKADLEKKISGLK 1133
Qy 268 EDNFEELQYFPNFYALLKSID-----SEKAEKETLITIMKTILIDFV-KMMVKYGTISP 321
Db 1134 QD-YEDLEDDKN-----KIEGLRNQRKIKELDDITGADVSVYLOKQKEEY----- 1181
Qy 322 EEGVSYLENLEDEMIALQTKNLEKNATDNISKLPAPSEKSHSEETDSTKEBAK-----ME 377
Db 1182 ESQIAKMQEKEKALGNDVKNK-EKTIKE--KELEIQSLQEKLDTEVEKEDAEKKKKEIE 1238
Qy 378 KEYGSLKDTKDDNSNPGKTDPEPKGTAYLEAIRKNIEWLKHDKGNKEDYDL-SKM 436
Db 1239 KEMKALQEE--KENVESSKNSTKDKKKLEDNLDKDTOKKLDMDTADNEKLLAKAKDLBAQL 1297
```



```

Db 207 HKVEENKSDHVKVEENKSDHKKIEVKKVEHEDEEDKKEKSEKKNKDNKEND 266
QY 141 LDGTPLAEDIVHKIAAIIYEENDRAVDFKIVSKLNLGLITSEQHTLEDVAEVLQKL 200
Db 267 EDNDEISDEVEDDDVEEDKQNDIDDDK-----KETDKHLEEBENBIEKE 315
QY 201 IS-KBANNYBEDPKPTSW-TENQAGKIPKVTTPMAAIQDGLAGNDETQVNTLTLNG 258
Db 316 FSDKKKNGKNDTKKEKSKDTEKESKDIEK-----EKSQKEKESK 357
QY 259 LERTKTYSEDNFERLOYPNFPVALLKSDSKSEKAKETLITIMKTLLIDFVKMMVYGT 318
Db 358 -KDEKEKKGKKEKE-----KSKDIEKE-KKQKQIEKSK-DTAKEKEDKD 403
QY 319 ISPEEGSVLENDEMIALQTKNLEKKNATDNISKLPAPSPKSH-----E 364
Db 404 IEKEKS-----KDEKLNKQNDKEDKDDNEKK-----KNDKQDHDNDNDMEIE 452
QY 365 ETDSTKEBAKME-----KEYGSLKDKSTK-DDNSNPGGKTDBPKGTEAYLEAI 412
Db 453 ENDEDEDEDEDEKKNKKGNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGN 512
QY 413 RNIEWLKKHDKKNGKEDYDLKMRDP-----INKQADAVVEKGILD 454
Db 513 NENNGNENENKENEKKNKIEIENVTVNANKENYEKINKNSEITITKSNID 564

```

RESULT 12

```

Q07380
ID Q07380 PRELIMINARY; PRT; 1790 AA.
AC Q07380; P89892;
DT 01-NOV-1996 (TRENBLrel. 01, Created)
DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)
DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical 206.5 kDa protein YDL058W.
GN US01 OR YDL058W.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RA Bloecker H., Brandt P.;
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z74106; CA98621.1; -.
DR EMBL; Z74105; CA98620.1; -.
DR SGD; S0002216; US01.
DR InterPro; IPR002017; Spectrin.
DR InterPro; IPR006955; Usol_p115_C.
DR Pfam; PF04871; Usol_p115_C; 1.
DR Pfam; PF04869; Usol_p115_head; 1.
SQ SEQUENCE 1790 AA; 206450 MW; 90062544F55A52EE CRC64;

```

```

Query Match
Best Local Similarity 20.3%; Score 183.5; DB 3; Length 1790;
Matches 120; Conservative 91; Mismatches 200; Indels 179; Gaps 23;

QY 29 KSLHN--RELSAERPLNEQIAEAEEDKIKKTPPENKPGQSNYSFVDNLLKAI-----T 82
Db 965 KSLANNYKQMAE---NESLIKAVEE-----SKNESSQLSNLKNQKIDMSQ 1008
QY 83 EKEIEKERQSTRSSPLDNKLNVEDVSTKNRKL I-----DDYSTKGLDHKQD--- 133
Db 1009 EKENFOIRGSIEKTEQLKKTISLEQTKHEIIKSDSSKDEYSQISLKEKLETATT 1068
QY 134 -DPDGLHQDGTPLTADIVHKIAAIIYEENDRAVDFKIVSKLNLGLITSEQAH----- 187
Db 1069 ANDENVNKISLTKTREETEELEALAA--YKNLKNLETKLETSEKALKEVKNESHKKEK 1126
QY 188 -TLEDEVAEVLQKLISKSEAN-----NVEED-PNKPTSWTENQAKIPEK 229

```

```

Db 1127 IQLEKEATETKQOLNSLRANLESLEKEHEDLAAQLKKYBEQIANKEROYNE-EISQLNDE 1185
QY 230 VTPMAAIQDGLAGNDETQVNTLTLTLNGLERT-----KTYSEDN--- 270
Db 1186 ITSTQENESIKK-KNDELEGEVKAAMKSTSEQSNLKKSEIDALNLOIKELKKKNTNEA 1244
QY 271 -----FEELYPPNF-----YALLKSIDSEKAK 294
Db 1245 SLLESIKSVESVETVKIKELQDECNFKVEVSELEDKASEDKNSKYLEQK-ESEKIKE 1303
QY 295 EKETLITIMKTLLIDFVKMMVK----- 315
Db 1304 ELDAKTEIKIQLKIEKITNLSKAKESSELSRLKKTSEERKNABEQLEKLNKNEIQIKNQ 1363
QY 316 -----YGTISPE--EGVSYLENDEMIALQTKNLEKKNATDNISKLPAPSEK 361
Db 1364 APEKERKLLNEGSSITIQEYSEKINTLE--DELIRLQNEHELKAKEDINTSELEKVSLS 1421
QY 362 SHEETDSTKEBAKMEKEVGSUKDS-TKDDNSNPGGKTDEP-----KGKTEAYLEAIRK 414
Db 1422 NDLLEEKQNTIKSLQDEILSYDKITRNDKLLSIERDKRDLSEKLEQLRAAQESKAK 1481
QY 415 NIEWLKKHDKKNGKEDYDLKMRDPINKQADAVVEKGILD-KEEAETIKR 463
Db 1482 VEEGLKUEEESSEKAELEKSEKMM-KKLESTIESNETELKSSMETIRK 1530

```

RESULT 13

```

Q815X5
ID Q815X5 PRELIMINARY; PRT; 2612 AA.
AC Q815X5;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-WAR-2003 (TRENBLrel. 23, Last annotation update)
DE Hypothetical protein.
GN PFL0350C.
OS Plasmodium falciparum (isolate 3D7).
OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
OX NCBI_TaxID=36329;
RN [1]
RP SEQUENCE FROM N.A.
RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
RA Pettea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,
RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
RA Fraser C.M., Barrell B.,
RT "Genome sequence of the human malaria parasite Plasmodium
RT falciparum."
RL Nature 419:498-511(2002).
DR EMBL; AE014845; AAN36159.1; -.
KW Hypothetical protein.
SQ SEQUENCE 2612 AA; 311421 MW; C950B6E466E3D1C5 CRC64;

```

```

Query Match
Best Local Similarity 22.3%; Score 183.5; DB 5; Length 2612;
Matches 114; Conservative 98; Mismatches 196; Indels 103; Gaps 23;

QY 27 QDKSLHNRELSAERP---LNEQIAEAEEDKIKKTPPENKPGQSNYSFVDNLLKAI 83
Db 991 REKEIDNERRNIEKENFLNHTYKKELEEKL-----NNKKEKNELEMLK--KE 1036
QY 84 KEKIEKERQSTRSSPLDNKLNVE-----DVDSTKNR--KLIDDYSTKSG 126
Db 1037 LESLEKEKEKQIIECEYNNLQNKKEEIQERNNLKQKELNSRMDRYNELLEDLKKRKE 1096
QY 127 LD---HKFODDPDG-----LHOLDGTPLTADIVHKIAAIIYEENDRAVDFKIVSKLNL 178

```


Db 1097 LEMKOKLLDDVQERRIKFLNEKNMOKKEKENEINYMKEELRKRILMIEVEVKVMIML 1156

Qy 179 GLI-----TSQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPE 228

Db 1157 EDIEKNKEKMIQVKEKENEKLEKTEKERNMI-----QNLEEKKEFKPLYLEQKYKENFE 1212

Qy 229 KVTMPAAIQQDGLAK--GENDETSTNTLTNGLERTKTYSEDNFEELQ--YFNF----- 280

Db 1213 NE-----KSLGAKFDENELQNEI-----GNEKKLHKERDNPFEOKKYBEEFRNKC 1262

Qy 281 ----YALLKSID--SEKAKEKETLITIMKTLIDFVGMVYKGTISPPEGVSYL----- 328

Db 1263 EKYBESIQKYDVLDEEDKNKYLIIKEQBELENVKNKI--YLDIEEKEKLYVQOEKLN 1320

Qy 329 ----ENL-----DEMIALQTN--KLEKNATDNISKL-----FPAPSEKSHETDSTKEEAAKME 377

Db 1321 LEKENLLVEKEQIEBELKNYKNFKEKEENDIKIKIINLSQQOEDLNKEKENIEKEQIE 1380

Qy 378 KEYGSLKSTKDDNSNPGKTDPEPKTEAYLEAIRKNIEWLKXDKK--GNKEDYDLSKM 436

Db 1381 KRKYDL--DEREESLNNDKAOMEESRKIFDQDLKIKKNKEELLNNDYKELKNREIHLMEKE 1439

Qy 437 RDFINKQADAYVEKILDKERAEAKRIYSS 467

Db 1440 KEYNKLSESNKKESLNSFVDEL--KEYSS 1468

RESULT 14

Q8IL45 PRELIMINARY; PRT; 3504 AA.

AC Q8IL45;

DT 01-MAR-2003 (TREMBLrel. 23, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein.

GN PF14.0404.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=36329;

RP SEQUENCE FROM N.A.

RC STRAIN=3D7;

RX MEDLINE=22255705; PubMed=12368864;

RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,

RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,

RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,

RA Chan M.-S., Nene V., Shallow J., Suh B., Peterson J., Angiuoli S.,

RA Perteau M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,

RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,

RA McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C.,

RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,

RA Fraser C.M., Barrett B.,

RT "Genome sequence of the human malaria parasite Plasmodium

RT falciparum.";

RL Nature 419:498-511 (2002).

DR EMBL; AE014822; AAN37017.1; -.

KW Hypothetical protein.

SQ SEQUENCE 3504 AA; 408303 MW; B8454D48D55BE4F0 CRC64;

Query Match 7.5%; Score 180.5; DB 5; Length 3504;

Best Local Similarity 20.5%; Pred. No. 1.7;

Matches 104; Conservative 93; Mismatches 172; Indels 139; Gaps 22;

Qy 32 HN-----RELSAE-----RPLNEQIAEAE-----EDKIKKTYPPENKPGQ----- 66

Db 982 HNLGKNEVDETMWEYIISPAKSELINEQETPPQVEVLSKNINNDKNGKRINNDOLE 1041

Qy 67 ----SNYSFVNLNLLKAITKEKIEKERQSGIRSPLDN-----KLNVEDVDSTNKLID 118

Db 1042 SPILSNEKKTDDIHI-----IEEKEIKD--PIHNDLENQNHKKHQD-----DKLE 1088

Qy 119 DYDSTKSLGDKHFQDDPGLHQLDGTPTABDIVHKIAARIYENDRADFVKIVSKLNL 178

Db 1089 EYSTLPNSKEMK-----NISGDNIENLLSNKKEENKHIHNIEDTTSK 1134

Qy 179 GLITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTEN-----OAGKI----- 226

Db 1135 DNEKNMQNYIHHDLG--KKDISRENTGDISPNKYLKNKENYGLDDKCKIKIYEEN 1191

Qy 227 PEKTPMAAIQDGLAKSENDET-----SNTLTNGLERTKTYSEDN 270

Db 1192 EYVTDVSGENKVSSEGNNDNIMPEYGNKNSQENSEDNIMTEKNGKNKSSEENTDDN 1251

Qy 271 FEELQYFPNFYALLKSIDSEKAKEKETLITIMKTLIDFVGMVYKGTISPPEGVSYL 330

Db 1252 ----YALLKSID--SEKAKEKETLITIMKTLIDFVGMVYKGTISPPEGVSYL----- 1285

Qy 331 LDEMIALQ----TRNKLEKNATDNISKLPAPSEKSHETDST-----KEEAAK----- 375

Db 1286 IDDNIMTEKNGKNNSQENSEDDNIMTEYKKNKSQENTDDNIMTEKNGKNKSSEENTDD 1345

Qy 376 --MEKEYGSLKSTKDDNSNPGKTDPEPKTEAYLEAIRKNIEWLKXDKKGNKEDYDL 433

Db 1346 NIMTEKNG--KDKSSEENANDNVMT--EKNGKNKSSEENADDNIMTEKNGKNKSSEENADD 1402

Qy 434 SKMRD--FINKQADAYVEKILDKERAE 459

Db 1403 NIIRENGKNKSQENSEDDNIMTEYGE 1430

RESULT 15

Q9JK25 PRELIMINARY; PRT; 1320 AA.

AC Q9JK25;

DT 01-OCT-2000 (TREMBLrel. 15, Created)

DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)

DT 01-OCT-2000 (TREMBLrel. 22, Last annotation update)

DE CLIP-170.

GN CYLN1.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RP SEQUENCE FROM N.A.

RC STRAIN=Wistar; TISSUE=Hippocampus;

RA Galjart N.;

RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RC STRAIN=Wistar; TISSUE=Hippocampus;

RX MEDLINE=21185938; PubMed=11290329;

RA Akhmanova A., Hoogenraad C.C., Drabek K., Stepanova T., Dortland B.,

RA Verkerk T., Vermeulen W., Burgering B.M., de Zeeuw C.I., Grosveld F.,

RA Galjart N.;

RT "CLASPs are CLIP-115 and -170 associating proteins involved in the

RT regional regulation of microtubule dynamics in motile fibroblasts.";

RL Cell 104:923-935 (2001).

DR EMBL; AJ237670; CAB92974.1; -.

DR InterPro; IPR000938; CAP-GLY.

DR InterPro; IPR001878; Znf.CCHC.

DR Pfam; PF01302; CAP.GLY.2.

DR SMART; SM00343; Znf.C2HC.1.

DR PROSITE; PS00845; CAP.GLY.1.2.

DR PROSITE; PS50245; CAP.GLY.2.2.

SQ SEQUENCE 1320 AA; 148264 MW; 87BE0F0463D20E9D CRC64;

Query Match 7.5%; Score 179.5; DB 11; Length 1320;

Best Local Similarity 22.6%; Pred. No. 0.62;

Matches 134; Conservative 77; Mismatches 211; Indels 171; Gaps 27;

Qy 22 KPGSQDQLSH--NRELSAERPLNEQIAEAEEDKIKKTYPPENKPGQSNYSFVNLNLLKA 80

Db 525 KPQGDVMSLSLQEIISA---LQEKLEVTTHDQNEVTSLKDHFTREEMFQKEIKALHA 581

Qy 81 ITEKEKIEKQSRSSPLD--NKLINVEDV----- 108

582	AT--EKLKSNESURSK-LDHANKENSVDJALWKSLETAIAHQOAMEELKVSFSGIG	638
109	-DS-----TKNRKLIIDDYSTKSGLDHFOODPDGLHQLODTPLTA-----EDI	151
639	TDASFAELKQIERLRDLYOHEITESLOSK-QDSERSAHAKEMESWAKMKIIEKEDS	697
152	VHKIAARIYENDRAV--FDKIVSKLNLNLGITESQHTLEDEVA-----EVLQKLISK	203
698	LEAVARLDTAEDQHLVEMEMELSKLQAEAIKKEKFPASAEAEAVSTQTSMQDVTNKLHQ	757
204	EANNYEEDPNKPTSWTENQACKIPEKVTTPMAA-----IQGLAGK--END--ETV--	249
758	-----EQFNLSSELE-----KURENLTDMEAKFKEDERDQLVAKAKELENDIAETMK	808
250	-----SNTFLTITNGLERRTKTYSEDNFEELQY-----PNFVALLKSI-----DSEK	291
809	MSGDNSSQLTKWMD-ELRLK--ERSVBELOLKTANENASILLQKSIQVTILKAOSQQ	864
292	EA--KEKETLITIMKTILDFVKWM-----VKYGTISPERSGVSY--LENLDEMI	335
865	EAAKHBEKELENKLELEKKMETSHYQCQDLAKAYEKASSETKIKHBEILQNFQKWL	924
336	-----ALOTKNKLEKNATDNISKL-----354	
925	VDTEDKLKAQEAERDLQMDEELKSQADKAAQTAEDAMQIMEQWTKETKTETLASLED	984
355	FPAPSEKSHBETOSTKEEAAKMEKEYGSLKSDTKODNSNPGGKTDEPKGKTEAYLEAIRK	414
985	TKQTNAKLQSELODTLKENNLKTVBELNKSKELLNEENQ-----KMEFPKKEITLQKAAQ	1040
415	NIEWLKHKDKGNKEDYDLKMRDFINKQADAYVEKGILDKKEEABAIKRIYVS	467
1041	KSOALSALQENVKLAEEGRTEDEVTSHQKLEERSVNLNQLLEMKKSLSPN	1093

Search completed: January 2, 2004, 12:35:46
Job time : 42 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 10:59:27 ; Search time 46 Seconds

(without alignments)
1614.870 Million cell updates/sec

Title: ABU66614

Perfect score: 2399

Sequence: 1 MGFLGTGTLVILVLPQAF.....EGILDKEAEAKRIYSSL 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_19Jun03.*

1: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.*
23: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.*
24: /SIDSI/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2399	100.0	468	21	AA18926
2	2399	100.0	468	22	AAU12216
3	2399	100.0	468	22	AA18926
4	2399	100.0	468	23	ABG95925
5	2399	100.0	468	24	ABU69097
6	2399	100.0	468	24	ABU69120
7	2399	100.0	468	24	ABU71580
8	2399	100.0	468	24	ABU72026
9	2399	100.0	468	24	ABU72183

10	2399	100.0	468	24	ABU66614	Human PRO polypept
11	2399	100.0	468	24	ABU66890	Human secreted/tra
12	2399	100.0	468	24	ABU59695	Novel secreted and
13	2399	100.0	468	24	ABG73287	Human neurosecret
14	2399	100.0	567	22	AAO13905	Human polypeptide
15	2392	99.7	468	22	AAW93910	Human polypeptide,
16	2384	99.4	468	21	AAW81955	Human HgII prote
17	2271	94.7	470	22	ABG12376	Novel human diagno
18	2102.5	87.6	471	23	ABW57046	Mouse ischaemic co
19	1974	82.3	396	24	AAO16322	Human secreted pro
20	594	24.8	316	22	ABG14006	Novel human diagno
21	475	19.8	212	22	ABG14003	Novel human diagno
22	184	7.7	665	21	AA18278	Plasmodium falcipa
23	180.5	7.5	1881	23	ABF73809	Candida albicans e
24	179	7.5	1558	21	AA181324	Plasmodium falcipa
25	179	7.5	1786	18	AAW24790	P. falciparum live
26	179	7.5	1787	23	AAU96699	Plasmodium falcipa
27	178.5	7.4	1392	20	AAW06999	Restin protein seq
28	175.5	7.3	1427	12	AA10534	Human 160kD mediat
29	174.5	7.3	1164	19	AAW40537	Group B streptococ
30	173	7.2	1099	19	AAW40538	Mutant C-beta prot
31	172.5	7.2	1164	17	AAW85781	Group B Streptococ
32	172.5	7.2	1164	21	AAW84459	Amino acid sequenc
33	172.5	7.2	1979	21	AA18171	Plasmodium falcipa
34	171	7.1	1093	19	AAW40540	Mutant C-beta prot
35	170	7.1	1213	22	AAW40016	Human polypeptide
36	169.5	7.1	1788	22	ABW62995	Drosophila melanog
37	169	7.0	1129	21	AAW84462	Amino acid sequenc
38	168	7.0	1164	19	AAW40541	Mutant C-beta prot
39	168	7.0	1164	21	AAW84463	Amino acid sequenc
40	167	7.0	493	13	AAW26944	P.falciparum USA 9
41	166.5	6.9	1135	21	AAW84460	Amino acid sequenc
42	166	6.9	1128	19	AAW40539	Mutant C-beta prot
43	166	6.9	1128	21	AAW84461	Amino acid sequenc
44	166	6.9	1192	21	AA18165	Plasmodium falcipa
45	166	6.9	2633	22	ABG06505	Novel human diagno

ALIGNMENTS

RESULT 1

AA18926

ID AA18926 standard; Protein; 468 AA.

XX

AC AA18926;

XX

DT 08-FEB-2001 (first entry)

XX

DE A novel polypeptide designated PRO5990.

XX

KW Secreted protein; transmembrane protein; PRO1484; PRO4334; PRO1122;

KW PRO1889; PRO1890; PRO1887; PRO1785; PRO4353; PRO4357; PRO4405; PRO4356;

KW PRO4352; PRO4380; PRO4354; PRO4408; PRO5737; PRO4425; PRO5990; PRO6030;

KW PRO4424; PRO4430; PRO4430; PRO4430; PRO4430; PRO4430; PRO4430; PRO4430;

KW insulinemia; kidney disorder; Bergers disease; nephropathy;

KW Schonlein-Henoch purpura; celiac disease; dermatitis herpetiformis;

KW Crohns disease.

XX Homo sapiens.

OS

XX

Key Location/Qualifiers

FT Modified-site 68..71 /note= "N-glycosylation site"

FT Modified-site 70..73 /note= "casein kinase II phosphorylation site"

FT Modified-site 82..85 /note= "casein kinase II phosphorylation site"

FT Modified-site 97..100 /note= "casein kinase II phosphorylation site"

FT Modified-site 125..128 /note= "casein kinase II phosphorylation site"

FT Modified-site 143..148 /note= "casein kinase II phosphorylation site"

CC polypeptide, H is a PRO20233 polypeptide and I is a PRO1890
 CC polypeptide. The sample comprises a cell suspected of expressing the A,
 CC B, C or D polypeptide. The A, B, C, D, E, F, G, H or I polypeptide is labeled with
 CC a detectable label or is attached to a solid support. The proteins are
 CC useful for linking a bioactive molecule to a cell expressing a
 CC polypeptide designated as A, B, C or D or E, F, G, H or I. The bioactive
 CC molecule is a toxin, a radiolabel or an antibody. The bioactive molecule
 CC causes death of the cell. A, B, C, D, E, F, G, H, or I, or antibodies
 CC against them are useful for modulating a biological activity of a cell
 CC expressing a polypeptide designated as A, B, C or D or E, F, G, H, or
 CC I. The cell is killed. The proteins are useful for identifying
 CC agonists or antagonists for the preparation of a medicament useful in
 CC the treatment of a condition which is responsive to the proteins, as
 CC molecular weight markers for protein electrophoresis purposes, and as
 CC therapeutic agents for treating sports-related joint problems,
 CC articular cartilage defects, osteoarthritis or rheumatoid arthritis.
 CC Nucleic acids encoding the proteins are useful as hybridisation probes,
 CC in chromosome and gene mapping, in the generation of anti-sense RNA and
 CC DNA, for the preparation of the proteins, to generate transgenic or
 CC knockout animals which are useful in the development and screening of
 CC therapeutic useful reagents, for chromosome identification, and in gene
 CC therapy. The antibody is useful as a therapeutic agent, in a diagnostic
 CC assay and for affinity purification of the protein from recombinant
 CC cell culture natural sources. The present sequence represents a novel
 CC secreted or transmembrane protein of the invention.
 XX
 SQ Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 23; Length 468;
 Best Local Similarity 100.0%; Pred. No. 4.1e-146;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGFLGTGTLVILVLPQAFKPGGQSKLHNRSLAERPLNEQIAEAEEDKIKTYPP 60
 DB 1 MGFLGTGTLVILVLPQAFKPGGQSKLHNRSLAERPLNEQIAEAEEDKIKTYPP 60

QY 61 ENKPGQSNYSFVDNLNLLKATEKEKIEKRSQSRSSPLDNVEDVDSTKNRKLIDDY 120
 DB 61 ENKPGQSNYSFVDNLNLLKATEKEKIEKRSQSRSSPLDNVEDVDSTKNRKLIDDY 120

QY 121 DSTKSGLDHKQDDPDGHLQDGTPLTAEDIVHKTAARIYEENDRAVEDKIVSKLLNLGL 180
 DB 121 DSTKSGLDHKQDDPDGHLQDGTPLTAEDIVHKTAARIYEENDRAVEDKIVSKLLNLGL 180

QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTEQAGKIPKVTMPMAAIOBGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTEQAGKIPKVTMPMAAIOBGL 240

QY 241 AKGENDETQSVNTLTITNGLRTKTYSDNFPEELQYFNFYALLKSIDSEKEAKEKETLI 300
 DB 241 AKGENDETQSVNTLTITNGLRTKTYSDNFPEELQYFNFYALLKSIDSEKEAKEKETLI 300

QY 301 TIMKTLDIFVKMVKYGTISPBEGVSYLENDEMIALQTKNLEKNATDNISKLFPAPSE 360
 DB 301 TIMKTLDIFVKMVKYGTISPBEGVSYLENDEMIALQTKNLEKNATDNISKLFPAPSE 360

QY 361 KSHETDSTKEAAKMEKEYGSLKSDTKDDNSNPGCKTDEPKGTEAYLEAIRKNIEWLK 420
 DB 361 KSHETDSTKEAAKMEKEYGSLKSDTKDDNSNPGCKTDEPKGTEAYLEAIRKNIEWLK 420

QY 421 KHKDKGNKEDYDLSKMRDPIFNKQADAYVEKGILDKKEAEAIKRIYSSL 468
 DB 421 KHKDKGNKEDYDLSKMRDPIFNKQADAYVEKGILDKKEAEAIKRIYSSL 468

RESULT 5
 ABU69097
 ID ABU69097 standard; Protein; 468 AA.
 AC ABU69097;
 XX
 XX
 DT 02-JUN-2003 (first entry)
 XX

DE Human PRO polypeptide #18.
 XX
 KW Human; PRO polypeptide; secreted and transmembrane protein; cancer;
 KW non-insulin dependent diabetes mellitus; septic shock; stroke;
 KW rheumatoid arthritis; graft-versus-host disease; cardiac ischaemia;
 KW psoriasis; inflammatory bowel disease; asthma; antidiabetic;
 KW cycostatic; immunosuppressive; antirheumatic; antiarthritic;
 KW cerebroprotective; vasotropic; antipsoriatic; antiinflammatory;
 KW antiasthmatic.
 OS Homo sapiens.
 XX
 FN US2003008348-A1.
 XX
 PD 09-JAN-2003.
 XX
 PF 26-DEC-2001; 2001US-0035855.
 XX
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-DEC-1999; 99WO-US28551.
 PR 22-DEC-1999; 99WO-US30720.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-JUN-2001; 2001WO-US17800.
 PR 29-JUN-2001; 2001WO-US19692.
 PR 09-JUL-2001; 2001WO-US21066.
 PR 15-MAY-1998; 98US-085579P.
 PR 15-DEC-1998; 98US-112514P.
 PR 22-DEC-1998; 98US-113300P.
 PR 23-DEC-1998; 98US-113430P.
 PR 23-DEC-1998; 98US-113605P.
 PR 22-JAN-1999; 99US-116843P.
 PR 23-MAR-1999; 99US-125774P.
 PR 23-MAR-1999; 99US-125778P.
 PR 24-MAR-1999; 99US-125826P.
 PR 31-MAR-1999; 99US-127035P.
 PR 05-APR-1999; 99US-127106P.
 PR 13-APR-1999; 99US-129122P.
 PR 21-APR-1999; 99US-130359P.
 PR 25-MAY-1999; 99US-135750P.
 PR 08-JUN-1999; 99US-138166P.
 PR 20-JUL-1999; 99US-144791P.
 PR 03-AUG-1999; 99US-146970P.
 PR 29-OCT-1999; 99US-162506P.
 PR 16-AUG-2001; 2001US-0931836.
 XX
 (GETH) GENENTECH INC.
 PA
 XX Desnoyers LA, Eaton DL, Goddard A, Godowski PJ, Gurney AL, Pan J;
 PI Stewart TA, Watanabe CK, Wood WI, Zhang Z;
 XX
 XX WPI; 2003-341326/32.
 DR N-PSDB; ACA06116.
 DR
 XX
 PT New PRO polypeptides and nucleic acid molecules, useful for diagnosing
 PT or treating diabetes mellitus, cancers, septic shock, inflammatory
 PT bowel disease or asthma, or in gene therapy, chromosome identification.
 PT or tissue typing
 XX
 PS Claim 12; Fig 36; 196pp; English.
 PS
 XX The present invention relates to the isolation of novel human PRO
 CC polypeptides, and the polynucleotide sequences encoding them. The
 CC PRO polypeptides are secreted and transmembrane proteins. The PRO
 CC polypeptides and polynucleotides are useful in diagnosing or
 CC treating non-insulin dependent diabetes mellitus, cancers, septic
 CC shock, rheumatoid arthritis, graft-versus-host disease, stroke, cardiac

CC	mellitus, hypo-insulinaemia), obesity, hyper-insulinaemia,
CC	haemoglobin-associated disorders (e.g. thalassemias), kidney disorders
CC	associated with decreased mesangial cell function (e.g. Berger disease),
CC	or other nephropathies associated with Schonlein-Henoch purpura,
CC	celliac disease, dermatitis herpetiformis or Crohn's disease. The PRO
CC	polynucleotide sequences may be used as hybridisation probes in
CC	chromosome and gene mapping, or in generating antisense RNA and DNA.
CC	They are also useful in preparing PRO polypeptides, in assays to
CC	identify other proteins or molecules involved in binding reaction, to
CC	generate transgenic animals or knockout animals, which in turn are
CC	useful in the development and screening of therapeutically useful
CC	reagents, for chromosome identification, and tissue typing. The PRO
CC	polypeptides and nucleic acid molecules are also useful in gene
CC	therapy, and as molecular weight markers for protein electrophoresis
CC	purposes. Anti-PRO antibodies may be used in diagnostic assays for
CC	PRO polypeptides, or for the affinity purification of the polypeptides
CC	from recombinant cell culture or natural sources. ABU69103-ABU69125
CC	represent the human PRO polypeptides of the invention.
XX	
SQ	Sequence 468 AA;
	Query Match 100.0%; Score 2399; DB 24; Length 468;
	Best Local Similarity 100.0%; Pred. No. 4.le-146;
	Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY	1 MGFLGTGTLVLVLPQAPFKPGSQKSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
DB	1 MGFLGTGTLVLVLPQAPFKPGSQKSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
QY	61 ENKPGQSNYSFVDNLNLLKAITEKEIKERQSISSPLDNKLVNVEDVSTNKRKLIDDY 120
DB	61 ENKPGQSNYSFVDNLNLLKAITEKEIKERQSISSPLDNKLVNVEDVSTNKRKLIDDY 120
QY	121 DSTKSLGDKHFDQDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDKIVSKLLNLGL 180
DB	121 DSTKSLGDKHFDQDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDKIVSKLLNLGL 180
QY	181 ITESQAHTLEDAVAVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTFMAAIQDGL 240
DB	181 ITESQAHTLEDAVAVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTFMAAIQDGL 240
QY	241 AKGENDETIVSNLTITUNGRLERTKTYSDNFEELQYFNFYALLKSIDSEKEAKEFTLI 300
DB	241 AKGENDETIVSNLTITUNGRLERTKTYSDNFEELQYFNFYALLKSIDSEKEAKEFTLI 300
QY	301 TIMKTLIDFVKMWVYGTISPPEGVSYLENLDEMIALQTKNKLKQATDNISKLPAPSE 360
DB	301 TIMKTLIDFVKMWVYGTISPPEGVSYLENLDEMIALQTKNKLKQATDNISKLPAPSE 360
QY	361 KSHEETDSTKEEAAKWEYGLSKDSTKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
DB	361 KSHEETDSTKEEAAKWEYGLSKDSTKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
QY	421 KHKKGKNEKEDYLSKQRFINKQADAYVEKGIIDKEEAEAKRIYSSL 468
DB	421 KHKKGKNEKEDYLSKQRFINKQADAYVEKGIIDKEEAEAKRIYSSL 468
RESULT 7	
ABU71580	
ID	ABU71580 standard; Protein; 468 AA.
XX	
AC	ABU71580;
XX	
DT	10-JUN-2003 (first entry)
XX	
DE	Human secreted polypeptide PRO5990.
XX	
KW	Human; gene therapy; tumour; cancer.
XX	
OS	Homo sapiens.
XX	
PN	US2003013855-A1.

XX	16-JAN-2003.
PD	
XX	03-MAY-2002; 2002US-0063616.
PF	
XX	30-DEC-1998; 98KR-0062142.
PR	08-MAR-1999; 99WO-US05028.
PR	14-MAY-1999; 99WO-US10733.
PR	30-DEC-1999; 99WO-US31274.
PR	18-FEB-2000; 2000WO-US04341.
PR	01-MAR-2000; 2000WO-US05601.
PR	02-MAR-2000; 2000WO-US05841.
PR	21-MAR-2000; 2000WO-US07532.
PR	22-MAY-2000; 2000WO-US14042.
PR	02-JUN-2000; 2000WO-US15264.
PR	24-AUG-2000; 2000WO-US23328.
PR	10-NOV-2000; 2000WO-US30873.
PR	01-DEC-2000; 2000WO-US32678.
PR	20-DEC-2000; 2000WO-US34956.
PR	28-FEB-2001; 2001WO-US06520.
PR	01-JUN-2001; 2001WO-US17800.
PR	14-MAY-1999; 99US-0311832.
PR	25-AUG-1999; 99US-0380137.
PR	25-AUG-1999; 99US-0380138.
PR	25-AUG-1999; 99US-0380139.
PR	15-SEP-1999; 99US-0380142.
PR	18-OCT-1999; 99US-0403297.
PR	12-NOV-1999; 99US-0423844.
PR	22-AUG-2000; 2000US-0644848.
PR	18-SEP-2000; 2000US-0646610.
PR	18-SEP-2000; 2000US-0665350.
PR	08-NOV-2000; 2000US-0709238.
PR	20-DEC-2000; 2000US-0747259.
PR	22-MAR-2001; 2001US-0816744.
PR	10-MAY-2001; 2001US-0854208.
PR	10-MAY-2001; 2001US-0854280.
PR	30-MAY-2001; 2001US-0870574.
PR	05-JUN-2001; 2001US-0874503.
PR	29-JUN-2001; 2001US-0869599.
PR	18-JUL-2001; 2001US-0908827.
PR	06-DEC-2001; 2001US-0006867.
XX	
PA	(GETH) GENENTECH INC.
XX	
XX	Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI	Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
XX	
DR	WPI; 2003-330485/31.
DR	N-PSDB; ACA58884.
XX	
PT	New isolated antibody specifically binding a PRO polypeptide, useful
PT	for the preparation of a medicament for treating disorders with the
PT	aberrant expression or activity of the PRO polypeptide, such as tumor
PT	conditions and cancer
XX	
PS	Example 17; Page 218-219; 406pp; English.
XX	
CC	The invention relates to an antibody that binds to a polypeptide with a
CC	fully defined sequence given in the specification. The methods and
CC	compositions (containing antibodies that specifically bind a PRO
CC	polypeptide) of the present invention are useful for the preparation of a
CC	medicament for the treatment of disorders associated with the aberrant
CC	expression or activity of the PRO polypeptide, such as tumour conditions
CC	and cancer. They can also be used to generate transgenic or knockout
CC	animals useful in the development and screening of therapeutically useful
CC	reagents. The PRO polypeptides and encoding nucleic acids can be used as
CC	molecular weight markers for protein electrophoresis, chromosome
CC	identification and tissue typing. The PRO polypeptides are useful to
CC	induce angiogenesis e.g wound healing; in the treatment of sports-related
CC	joint problems, articular cartilage defects, osteoarthritis or rheumatoid
CC	arthritis; diabetes; hyperinsulinaemia and hypoinsulinaemia. The
CC	antibodies may be used in various diagnostic, competitive binding and/or

DE Human PRO polypeptide #75.
XX Human; PRO polypeptide; secreted and transmembrane protein;
KW anti-PRO antibody; diagnostic assay; gene expression.
XX
OS Homo sapiens.
XX
FN US2003023042-A1.
XX
PD 30-JAN-2003.
XX
PF 01-MAY-2002; 2002US-0063502.
XX
PR 06-DEC-2001; 2001US-0006867.
XX
XX (GETH) GENENTECH INC.
PA Eaton DL, Filvaroff E, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Watanabe CK, Wood WI;
XX WPI; 2003-331484/31.
DR N-PSDB; ACA63447.
XX
XX Novel monoclonal antibody that binds to secreted and transmembrane
PT polypeptide, useful for detecting and purifying the polypeptide and
PT also for treating conditions responsive to the antibody -
XX
PS Disclosure; Fig 150; 408pp; English.
XX
XX The present invention relates to the isolation of novel human PRO
CC polypeptides, and the polynucleotide sequences encoding them. The
CC PRO polypeptides are secreted and transmembrane proteins. The PRO
CC polypeptides and polynucleotides are useful for preparing a
CC medicament useful in the treatment of a condition responsive to
CC anti-PRO antibody. Anti-PRO antibodies are useful in diagnostic
CC assays for PRO, by detecting its expression in specific cells,
CC tissues or serum, and for affinity purification of PRO from
CC recombinant cell culture or natural sources. ABU72109-ABU72192
CC represent the human PRO polypeptides of the invention.
XX
XX Sequence 468 AA;
Query Match 100.0%; Score 2399; DB 24; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGWIILVLPICQAFKPGSGQSKLHNRELSAERPLNEQIAEAEEDKIKTYPP 60
DB 1 MGFLGTGWIILVLPICQAFKPGSGQSKLHNRELSAERPLNEQIAEAEEDKIKTYPP 60
QY 61 ENKPGQSNYSFVDNLNLKKAITEKEIKERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
DB 61 ENKPGQSNYSFVDNLNLKKAITEKEIKERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
QY 121 DSTKSGLDHKEQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180
DB 121 DSTKSGLDHKEQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180
QY 181 ITESQAHTLEDAEVAEVLQKLSKEANNYEENPNKPTSWTENQAGKIPKVTMPMAIQDGL 240
DB 181 ITESQAHTLEDAEVAEVLQKLSKEANNYEENPNKPTSWTENQAGKIPKVTMPMAIQDGL 240
QY 241 AKGENDETVSNLTLTNGLERTKTYSDNPFEEQLQYPPNFVALLKSIDSEKEAKEKETLI 300
DB 241 AKGENDETVSNLTLTNGLERTKTYSDNPFEEQLQYPPNFVALLKSIDSEKEAKEKETLI 300
QY 301 TIMKTLIDFVKMWKYGTISPEEGVSYLENLDEMIALQTKNLEKNATDNISKLFPAPSE 360
DB 301 TIMKTLIDFVKMWKYGTISPEEGVSYLENLDEMIALQTKNLEKNATDNISKLFPAPSE 360
QY 361 KSHETDSTKEAAKMEKEYSGLKSDTKDNSNPGGKTDEPKGTEYLAIRKNIEWLK 420
DB 361 KSHETDSTKEAAKMEKEYSGLKSDTKDNSNPGGKTDEPKGTEYLAIRKNIEWLK 420

QY 421 KHKDKGNKEDYDLKSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
DB 421 KHKDKGNKEDYDLKSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
RESULT 10
ABU66614
ID ABU66614 standard; Protein; 468 AA.
XX
AC ABU66614;
XX
DT 23-MAY-2003 (first entry)
XX
DE Human PRO polypeptide #45.
XX
KW Human; PRO polypeptide; secreted and transmembrane protein;
KW tumour necrosis factor-alpha; TNF-alpha; blood; proliferation;
KW differentiation; chondrocyte; tumour; genetic disorder;
KW cytostatic.
XX
OS Homo sapiens.
XX
FN US2003036180-A1.
XX
PD 20-FEB-2003.
XX
PF 09-MAY-2002; 2002US-0143114.
XX
PR 31-MAR-1997; 97WO-US05230.
PR 12-JUN-1998; 98WO-US12456.
PR 14-JUL-1998; 98WO-US14552.
PR 28-AUG-1998; 98WO-US17888.
PR 10-SEP-1998; 98WO-US18824.
PR 14-SEP-1998; 98WO-US19093.
PR 14-SEP-1998; 98WO-US19094.
PR 14-SEP-1998; 98WO-US19177.
PR 16-SEP-1998; 98WO-US19330.
PR 17-SEP-1998; 98WO-US19437.
PR 07-OCT-1998; 98WO-US21141.
PR 29-OCT-1998; 98WO-US22991.
PR 29-OCT-1998; 98WO-US22992.
PR 20-NOV-1998; 98WO-US24855.
PR 01-DEC-1998; 98WO-US25108.
PR 05-JAN-1999; 99WO-US00106.
PR 08-MAR-1999; 99WO-US05028.
PR 10-MAR-1999; 99WO-US05190.
PR 20-APR-1999; 99WO-US08615.
PR 14-MAY-1999; 99WO-US10733.
PR 02-JUN-1999; 99WO-US12252.
PR 01-SEP-1999; 99WO-US20411.
PR 08-SEP-1999; 99WO-US20594.
PR 13-SEP-1999; 99WO-US20944.
PR 13-SEP-1999; 99WO-US21090.
PR 15-SEP-1999; 99WO-US21547.
PR 05-OCT-1999; 99WO-US23089.
PR 29-NOV-1999; 99WO-US28214.
PR 30-NOV-1999; 99WO-US28313.
PR 30-NOV-1999; 99WO-US28409.
PR 01-DEC-1999; 99WO-US28301.
PR 01-DEC-1999; 99WO-US28634.
PR 02-DEC-1999; 99WO-US28551.
PR 02-DEC-1999; 99WO-US28564.
PR 02-DEC-1999; 99WO-US28565.
PR 16-DEC-1999; 99WO-US30095.
PR 20-DEC-1999; 99WO-US30911.
PR 20-DEC-1999; 99WO-US30999.
PR 22-DEC-1999; 99WO-US30720.
PR 30-DEC-1999; 99WO-US31243.
PR 30-DEC-1999; 99WO-US31274.
PR 05-JAN-2000; 2000WO-US00219.
PR 06-JAN-2000; 2000WO-US00277.
PR 06-JAN-2000; 2000WO-US00376.

XX OS Homo sapiens.
 XX XX US2003032155-A1.
 XX PN 13-FEB-2003.
 XX PD
 XX PF
 XX PF 03-MAY-2002; 2002US-0137865.
 XX PR 31-MAR-1997; 97WO-US05230.
 PR 12-JUN-1998; 98WO-US12456.
 PR 14-JUL-1998; 98WO-US14552.
 PR 28-AUG-1998; 98WO-US17888.
 PR 10-SEP-1998; 98WO-US18824.
 PR 14-SEP-1998; 98WO-US19093.
 PR 14-SEP-1998; 98WO-US19094.
 PR 14-SEP-1998; 98WO-US19177.
 PR 16-SEP-1998; 98WO-US19330.
 PR 17-SEP-1998; 98WO-US19437.
 PR 07-OCT-1998; 98WO-US21141.
 PR 29-OCT-1998; 98WO-US22991.
 PR 29-OCT-1998; 98WO-US22992.
 PR 20-NOV-1998; 98WO-US24855.
 PR 01-DEC-1998; 98WO-US25108.
 PR 05-JAN-1999; 99WO-US00106.
 PR 08-MAR-1999; 99WO-US05028.
 PR 10-MAR-1999; 99WO-US05190.
 PR 20-APR-1999; 99WO-US08615.
 PR 14-MAY-1999; 99WO-US10733.
 PR 02-JUN-1999; 99WO-US12252.
 PR 01-SEP-1999; 99WO-US20111.
 PR 08-SEP-1999; 99WO-US20594.
 PR 13-SEP-1999; 99WO-US20944.
 PR 15-SEP-1999; 99WO-US21090.
 PR 15-SEP-1999; 99WO-US21547.
 PR 05-OCT-1999; 99WO-US23089.
 PR 29-NOV-1999; 99WO-US28214.
 PR 30-NOV-1999; 99WO-US28313.
 PR 30-NOV-1999; 99WO-US28409.
 PR 01-DEC-1999; 99WO-US28301.
 PR 01-DEC-1999; 99WO-US28634.
 PR 02-DEC-1999; 99WO-US28551.
 PR 02-DEC-1999; 99WO-US28564.
 PR 02-DEC-1999; 99WO-US28565.
 PR 16-DEC-1999; 99WO-US30095.
 PR 20-DEC-1999; 99WO-US30911.
 PR 20-DEC-1999; 99WO-US30999.
 PR 22-DEC-1999; 99WO-US30720.
 PR 30-DEC-1999; 99WO-US31243.
 PR 30-DEC-1999; 99WO-US31274.
 PR 05-JAN-2000; 2000WO-US00219.
 PR 06-JAN-2000; 2000WO-US00277.
 PR 11-FEB-2000; 2000WO-US00376.
 PR 18-FEB-2000; 2000WO-US04341.
 PR 18-FEB-2000; 2000WO-US04342.
 PR 22-FEB-2000; 2000WO-US04414.
 PR 24-FEB-2000; 2000WO-US04914.
 PR 24-FEB-2000; 2000WO-US05004.
 PR 01-MAR-2000; 2000WO-US05601.
 PR 02-MAR-2000; 2000WO-US05746.
 PR 02-MAR-2000; 2000WO-US05841.
 PR 10-MAR-2000; 2000WO-US06319.
 PR 15-MAR-2000; 2000WO-US06884.
 PR 20-MAR-2000; 2000WO-US07377.
 PR 21-MAR-2000; 2000WO-US07532.
 PR 30-MAR-2000; 2000WO-US08439.
 PR 17-MAY-2000; 2000WO-US13705.
 PR 22-MAY-2000; 2000WO-US14042.
 PR 30-MAY-2000; 2000WO-US14941.
 PR 02-JUN-2000; 2000WO-US15264.
 PR 28-JUL-2000; 2000WO-US20710.
 PR 11-AUG-2000; 2000WO-US22031.

PR 23-AUG-2000; 2000WO-US23522.
 PR 24-AUG-2000; 2000WO-US23328.
 PR 08-NOV-2000; 2000WO-US30952.
 PR 10-NOV-2000; 2000WO-US30873.
 PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-MAR-2001; 2001WO-US06666.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 22-JUN-2001; 2001WO-US19692.
 PR 29-JUN-2001; 2001WO-US20116.
 PR 09-JUL-2001; 2001WO-US21086.
 PR 20-DEC-2000; 2000WO-US21735.
 PR 28-FEB-2001; 2001US-0796498.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0808689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 18-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 25-MAY-2001; 2001US-0866034.
 PR 01-JUN-2001; 2001US-0872035.
 PR 05-JUN-2001; 2001US-0874503.
 PR 14-JUN-2001; 2001US-0882636.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.

(GETH) GENENTECH INC.

Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W, Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S, Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z; WPI; 2003-331925/31.
 N-PSDB; ACA04068.

New secreted and transmembrane nucleic acids and polypeptides, designated as PRO, useful for treating inflammation, organ failure, atherosclerosis, cardiac injury, infertility, birth defects, premature aging, AIDS, or cancer

Claim 12; Fig 90; 659pp; English.

The invention relates to an isolated nucleic acid comprising, or which is at least 80% identical to, or the full-length coding sequence of, any of the 275 nucleotide sequences, encoding the corresponding PRO polypeptide (one of 275 secreted or transmembrane proteins). The nucleic acid further comprises the full-length coding sequence of the DNA deposited under American Type Culture Collection (ATCC) accession number in a list given in the specification. Also included are vectors and host cells for producing PRO proteins, PRO fusion proteins, anti-PRO antibodies, PRO extracellular domains and mature sequences, methods of detecting PRO proteins, methods for stimulating the release of TNF-alpha (tumour necrosis factor alpha) from human blood, and the proliferation of differentiation of chondrocyte cells, the proliferation of, or gene expression in pericyte cells, the release or proteoglycans from cartilage, proliferation of inner ear utricular supporting cells, the proliferation of T-lymphocyte cells, the release of a cytokine from peripheral blood mononuclear cells (PBMC), or the proliferation of endothelial cells), a method for modulating the uptake of glucose or free fatty acid (FFA) by skeletal muscle cells, a method for inhibiting the binding of A-peptide to factor VIIA, or the differentiation of adipocyte cells, a method for detecting the presence of a tumour in a mammal and an oligonucleotide probe derived

CC from any of the nucleotide sequences cited above. The nucleic acids and
CC polypeptides are useful for treating inflammatory diseases, organ
CC failure, atherosclerosis, cardiac injury, infertility, birth defects,
CC premature aging, AIDS (acquired immunodeficiency syndrome), cancer, or
CC diabetic complications. The nucleic acids are useful as hybridisation
CC probes, in chromosome and gene mapping, and in generating antisense RNA
CC or DNA. The polypeptides are useful as pharmaceuticals, diagnostics,
CC biosensors or bioreactors. Both are useful in tissue typing.
XX The present sequence represents a PRO protein of the invention.
XX
SQ Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 24; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGWIILVLPQAFKPGSQSKLHNRLSARPLNEQIAEAEEDKIKTYPP 60
DB 1 MGFLGTGWIILVLPQAFKPGSQSKLHNRLSARPLNEQIAEAEEDKIKTYPP 60
QY 61 ENKFGQSNYSFVDNLNLKAI TEKEIKERQSRSSPLDNKLNVEDVDSTNKRKLIDDY 120
DB 61 ENKFGQSNYSFVDNLNLKAI TEKEIKERQSRSSPLDNKLNVEDVDSTNKRKLIDDY 120
QY 121 DSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVPDKIVSKLLNLGL 180
DB 121 DSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVPDKIVSKLLNLGL 180
QY 181 ITESQAHLEDEVAEVLQKLSKANNYEEDPNKPTSWTENQAGKIPKVTMPMAIQDGL 240
DB 181 ITESQAHLEDEVAEVLQKLSKANNYEEDPNKPTSWTENQAGKIPKVTMPMAIQDGL 240
QY 241 AKGNDETIVSNLTITNGLERRTKTSYSDNPEELQYFPNFYALLKSIDSEKEAKEKETLI 300
DB 241 AKGNDETIVSNLTITNGLERRTKTSYSDNPEELQYFPNFYALLKSIDSEKEAKEKETLI 300
QY 301 TIMKTLIDFVQWVKYGTISPEEGVSVLENDEMIALQTKNLEKATDNISKLPAPSE 360
DB 301 TIMKTLIDFVQWVKYGTISPEEGVSVLENDEMIALQTKNLEKATDNISKLPAPSE 360
QY 361 KSHETSTKEEAKMEKYGSLKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
DB 361 KSHETSTKEEAKMEKYGSLKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLK 420
QY 421 KHKDKGNKEDYDLKMRDFINKQADAYVEKGILDKKEAAEAIKRIYSSL 468
DB 421 KHKDKGNKEDYDLKMRDFINKQADAYVEKGILDKKEAAEAIKRIYSSL 468

RESULT 12
ABU59695
ID ABU59695 standard; Protein; 468 AA.
XX
XX AC ABU59695;
XX
XX

DT 13-MAY-2003 (first entry)

DE Novel secreted and transmembrane protein PRO5990.

XX Human; PRO; hypertrophy of neonatal heart; angiogenesis; wound healing;
XX cardiac insufficiency disorder; cancer; tumour; immune response;
XX adrenal cortical capillary endothelial growth; c-fos induction;
XX vascular endothelial growth factor inhibition; VEGF inhibition;
XX endothelial cell growth inhibitor; T-lymphocytes stimulation;
XX retinal neurons cell survival; rod photoreceptor cell survival;
XX retinal disorder; retinitis pigmentosa; kidney disorder;
XX mammalian kidney mesangial cell proliferation; Berger disease;
XX dermatitis; herpeticiformis; Crohn's disease; chondrocyte proliferation;
XX chondrocyte redifferentiation; sports injury; arthritis.

OS Homo sapiens.

XX US2003017563-A1.

PN

XX 23-JAN-2003.
PD 07-MAY-2002; 2002US-0140808.
XX
XX
XX 31-MAR-1997; 97WO-US05230.
XX 12-JUN-1998; 98WO-US12456.
XX 14-JUL-1998; 98WO-US14552.
XX 28-AUG-1998; 98WO-US17888.
XX 10-SEP-1998; 98WO-US18824.
XX 14-SEP-1998; 98WO-US19093.
XX 14-SEP-1998; 98WO-US19094.
XX 14-SEP-1998; 98WO-US19177.
XX 16-SEP-1998; 98WO-US19330.
XX 17-SEP-1998; 98WO-US19437.
XX 07-OCT-1998; 98WO-US21141.
XX 29-OCT-1998; 98WO-US22991.
XX 29-OCT-1998; 98WO-US22992.
XX 20-NOV-1998; 98WO-US24855.
XX 01-DEC-1998; 98WO-US25108.
XX 05-JAN-1999; 99WO-US00106.
XX 08-MAR-1999; 99WO-US05028.
XX 10-MAR-1999; 99WO-US05190.
XX 20-APR-1999; 99WO-US08615.
XX 14-MAY-1999; 99WO-US10733.
XX 02-JUN-1999; 99WO-US12252.
XX 01-SEP-1999; 99WO-US20111.
XX 08-SEP-1999; 99WO-US20594.
XX 13-SEP-1999; 99WO-US20944.
XX 15-SEP-1999; 99WO-US21090.
XX 15-SEP-1999; 99WO-US21547.
XX 05-OCT-1999; 99WO-US23089.
XX 29-NOV-1999; 99WO-US28214.
XX 30-NOV-1999; 99WO-US28313.
XX 30-NOV-1999; 99WO-US28409.
XX 01-DEC-1999; 99WO-US28301.
XX 01-DEC-1999; 99WO-US28634.
XX 02-DEC-1999; 99WO-US28551.
XX 02-DEC-1999; 99WO-US28564.
XX 02-DEC-1999; 99WO-US28565.
XX 16-DEC-1999; 99WO-US30095.
XX 20-DEC-1999; 99WO-US30911.
XX 20-DEC-1999; 99WO-US30999.
XX 22-DEC-1999; 99WO-US30720.
XX 30-DEC-1999; 99WO-US31243.
XX 30-DEC-1999; 99WO-US31274.
XX 05-JAN-2000; 2000WO-US00219.
XX 06-JAN-2000; 2000WO-US00277.
XX 06-JAN-2000; 2000WO-US00376.
XX 11-FEB-2000; 2000WO-US03565.
XX 18-FEB-2000; 2000WO-US04341.
XX 18-FEB-2000; 2000WO-US04342.
XX 22-FEB-2000; 2000WO-US04414.
XX 24-FEB-2000; 2000WO-US04914.
XX 01-MAR-2000; 2000WO-US05004.
XX 02-MAR-2000; 2000WO-US05601.
XX 02-MAR-2000; 2000WO-US05746.
XX 10-MAR-2000; 2000WO-US05841.
XX 15-MAR-2000; 2000WO-US06319.
XX 20-MAR-2000; 2000WO-US06884.
XX 21-MAR-2000; 2000WO-US07377.
XX 30-MAR-2000; 2000WO-US07532.
XX 30-MAR-2000; 2000WO-US08439.
XX 17-MAY-2000; 2000WO-US13705.
XX 22-MAY-2000; 2000WO-US14042.
XX 30-MAY-2000; 2000WO-US14941.
XX 02-JUN-2000; 2000WO-US15264.
XX 28-JUL-2000; 2000WO-US20710.
XX 11-AUG-2000; 2000WO-US22031.
XX 23-AUG-2000; 2000WO-US23522.
XX 24-AUG-2000; 2000WO-US23328.
XX 08-NOV-2000; 2000WO-US30952.
XX 10-NOV-2000; 2000WO-US30873.

PR 01-DEC-2000; 2000WO-US32678.
 PR 20-DEC-2000; 2000WO-US34956.
 PR 28-FEB-2001; 2001WO-US06520.
 PR 01-MAR-2001; 2001WO-US06656.
 PR 25-MAY-2001; 2001WO-US17092.
 PR 01-JUN-2001; 2001WO-US17800.
 PR 20-JUN-2001; 2001WO-US19692.
 PR 22-JUN-2001; 2001WO-US20116.
 PR 29-JUN-2001; 2001WO-US21066.
 PR 09-JUL-2001; 2001WO-US21735.
 PR 20-DEC-2000; 2000US-0747259.
 PR 28-FEB-2001; 2001US-0796498.
 PR 09-MAR-2001; 2001US-0802706.
 PR 14-MAR-2001; 2001US-0808689.
 PR 22-MAR-2001; 2001US-0816744.
 PR 05-APR-2001; 2001US-0828366.
 PR 10-MAY-2001; 2001US-0854208.
 PR 10-MAY-2001; 2001US-0854280.
 PR 18-MAY-2001; 2001US-0860216.
 PR 25-MAY-2001; 2001US-0866028.
 PR 25-MAY-2001; 2001US-0866034.
 PR 01-JUN-2001; 2001US-0872035.
 PR 05-JUN-2001; 2001US-0874503.
 PR 14-JUN-2001; 2001US-0882636.
 PR 19-JUN-2001; 2001US-0886342.
 PR 21-JUN-2001; 2001US-0887879.
 PR 18-JUL-2001; 2001US-0908827.
 PR 06-AUG-2001; 2001US-0924419.
 PR 09-AUG-2001; 2001US-0927796.
 PR 16-AUG-2001; 2001US-0931836.
 PR 19-DEC-2001; 2001US-0028072.
 XX
 PA (GETH) GENENTECH INC.
 XX
 XX Baker KP, Beresini M, DeForge L, Desnoyers L, Filvaroff E, Gao W;
 PI Gerritsen ME, Goddard A, Godowski PJ, Gurney AL, Sherwood S;
 PI Smith V, Stewart TA, Tumas D, Watanabe CK, Wood WI, Zhang Z;
 XX
 DR WPI; 2003-148238/14.
 DR N-PSDB; ABX89185.
 XX
 PT Novel isolated PRO polypeptides e.g., PRO826, PRO1068, PRO1184, PRO1346
 PT and PRO1375, which stimulate proliferation of stimulated T-lymphocytes
 PT are therapeutically useful for enhancing immune response and in cancer
 PT treatments -
 XX
 PS Claim 12; Fig 90; 659pp; English.
 XX
 CC The invention describes an isolated human PRO polypeptide. The PRO
 CC polypeptides are useful in detecting PRO polypeptides in a sample, in
 CC linking a bioactive molecule to a cell expressing a PRO polypeptide, and
 CC in modulating at least one biological activity of a cell expressing a PRO
 CC polypeptide. PRO1312 stimulates hypertrophy of neonatal heart and is thus
 CC useful for treating cardiac insufficiency disorders. PRO1154 and PRO186
 CC stimulate adrenal cortical capillary endothelial growth, and PRO536,
 CC PRO943, PRO826, PRO1068 or PRO535, PRO826, PRO819, PRO1126,
 CC PRO1360 and PRO1387 induce c-fos in endothelial cells, and are thus
 CC useful for treating conditions or disorders where angiogenesis would be
 CC beneficial, e.g. wound healing and antagonist of this polypeptide are
 CC useful for treating cancerous tumours. PRO812 inhibits vascular
 CC endothelial growth factor (VEGF) stimulated proliferation of endothelial
 CC cells and is thus useful for inhibiting endothelial cell growth in
 CC mammals which would be beneficial in inhibiting tumour growth. PRO826,
 CC PRO1068, PRO1184, PRO1346 and PRO1375 stimulate proliferation of
 CC stimulated T-lymphocytes and are therapeutically useful for enhancing
 CC immune response. PRO828, PRO826, PRO1068 or PRO1132 enhance survival of
 CC retinal neurons cells (PRO1132 is also enhances survival/proliferation of
 CC rod photoreceptor cells) and therefore are useful for treating retinal
 CC disorders of injuries, e.g. retinitis pigmentosa, AMD. PRO819, PRO813
 CC and PRO1066 induce proliferation of mammalian kidney mesangial cells,
 CC and therefore are useful for treating kidney disorders associated with
 CC decreased mesangial cell function such as Berger disease or other
 CC nephropathies associated with dermatitis, herpeticiformis or Crohn's

CC disease. PRO1310, PRO844, PRO1312, PRO1192 and PRO1387 induce the
 CC proliferation and/or redifferentiation of chondrocytes in culture and
 CC are thus useful for treating sports injuries, and arthritis. This
 CC is the amino acid sequence of a novel human PRO protein.
 XX
 SQ Sequence 468 AA;
 Query Match 100.0%; Score 2399; DB 24; Length 468;
 Best Local Similarity 100.0%; Pred. No. 4, 1e-146;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGFLGTGTLVLVLPLOAPFKPGSGSQSLHNRLELSAERPLNEQIAEAEEDKIKTYPP 60
 DB 1 MGFLGTGTLVLVLPLOAPFKPGSGSQSLHNRLELSAERPLNEQIAEAEEDKIKTYPP 60
 QY 61 ENKPGQSNYSFVNDNLNLLKAITTEKEIEKERQSRSSPLDNKLNVEDVDSTKPKRLIDDY 120
 DB 61 ENKPGQSNYSFVNDNLNLLKAITTEKEIEKERQSRSSPLDNKLNVEDVDSTKPKRLIDDY 120
 QY 121 DSTKSGLDHFKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFDKIVSKLLNLGL 180
 DB 121 DSTKSGLDHFKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFDKIVSKLLNLGL 180
 QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPAAIQDGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPAAIQDGL 240
 QY 241 AKGENDETVNTLTITNGLERRTKTYSEDNFELOFPNFYALLKSIDSEKAKEKETLI 300
 DB 241 AKGENDETVNTLTITNGLERRTKTYSEDNFELOFPNFYALLKSIDSEKAKEKETLI 300
 QY 301 TIMKTLLIDFVMMVKYGTITSPREGVSYLENDEMIALQTKNKLKKNATDNISKLPAPSE 360
 DB 301 TIMKTLLIDFVMMVKYGTITSPREGVSYLENDEMIALQTKNKLKKNATDNISKLPAPSE 360
 QY 361 KSHEETDSTKEEAAKMEYGLSKDSTKDNNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
 DB 361 KSHEETDSTKEEAAKMEYGLSKDSTKDNNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
 QY 421 KHDKGNKEDYDLSKMRDFINKQADAVVEKGILDKEEAIAKRIYSSL 468
 DB 421 KHDKGNKEDYDLSKMRDFINKQADAVVEKGILDKEEAIAKRIYSSL 468
 RESULT 13
 ABG73287
 ID ABG73287 standard; Protein; 468 AA.
 XX
 AC ABG73287;
 XX
 DT 28-APR-2003 (first entry)
 XX
 DE Human neurosecretory protein-1 (HUNSP-1).
 XX
 KW Human; neurosecretory protein-1; HUNSP-1; neurological disorder;
 KW epilepsy; Huntington's disease; dementia; Alzheimer's disease;
 KW multiple sclerosis; Parkinson's disease; schizophrenia; amnesia;
 KW endocrine disorder; hypogonadism; Sheehan's syndrome; leukaemia;
 KW Hashimoto's disease; Grave's disease; Cushing's disease; cancer;
 KW adenocarcinoma; diabetes insipidus; melanoma; myeloma sarcoma;
 KW bone cancer; breast cancer; cancer of the pancreas; brain cancer;
 KW prostate cancer; skin cancer; thyroid cancer; antidiabetic; nootropic;
 KW neuroprotective; neuroleptic; cytostatic; anticonvulsant; thyromimetic;
 KW lymphoma; antiparkinsonian; antithyroid; immunosuppressive; antithyroid;
 KW pancreatic islet cell cDNA library; ISLNOT01; incyte clone 2379427.
 XX
 OS Homo sapiens.
 XX
 XX US2002156014-A1.
 PN
 XX
 PD 24-OCT-2002.
 XX
 XX 14-FEB-2002; 2002US-0077040.
 PF

XX 17-APR-1998; 98US-0062601.
XX (INCY-) INCYTE PHARM INC.
XX
XX Lal P, Corley NC, Patterson C;
XX WPI; 2003-238117/23.
XX N-PSDB; ABX11115.
XX
XX New human neurosecretory proteins (HUNSP) and polynucleotides, useful
XX for diagnosing, treating or preventing e.g. epilepsy, Alzheimer's
XX disease, schizophrenia, Grave's disease, leukaemia, lymphoma, or breast
XX or brain cancer -
XX
XX Claim 1; Fig 1; 45pp; English.
XX
XX The present invention relates to the isolation of human neurosecretory
XX proteins -1 and -2 (HUNSP-1 and HUNSP-2), and the polynucleotide
XX sequences encoding them. The HUNSP polypeptides, polynucleotides, and
XX agonists are useful for treating a disease or condition associated with
XX decreased expression of functional HUNSP. Antagonists of HUNSP are
XX useful for treating a disease or condition associated with
XX overexpression of functional HUNSP. HUNSP antibodies are useful for
XX diagnosing a condition or disease associated with the expression of
XX HUNSP. The sequences of the invention are particularly useful for
XX diagnosing, treating or preventing neurological disorders
XX (e.g. epilepsy, Huntington's disease, dementia, Alzheimer's disease,
XX multiple sclerosis, Parkinson's disease, schizophrenia, amnesia),
XX endocrine disorders (e.g. hypogonadism, Sheehan's syndrome, diabetes
XX insipidus, Hashimoto's disease, Grave's disease, Cushing's disease),
XX and cancers (e.g. adenocarcinoma, leukaemia, lymphoma, melanoma,
XX myeloma sarcoma, bone cancer, breast cancer, cancer of the pancreas,
XX brain cancer, prostate cancer, skin cancer or thyroid cancer).
XX The present sequence represents HUNSP-1. The cDNA encoding HUNSP-1
XX is isolated from pancreatic islet cell cDNA library (ISLNTOT01) Incyte
XX clone 2379427.
XX
XX Sequence 468 AA;

Query Match 100.0%; Score 2399; DB 24; Length 468;
Best Local Similarity 100.0%; Pred. No. 4.1e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTTWLVLPQAPFKPGGQDLSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTTWLVLPQAPFKPGGQDLSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVNDLNLKAI TEKEKERQSISSPLDNKLNVEDVDSTKNRKLIDDY 120
Db 61 ENKPGQSNYSFVNDLNLKAI TEKEKERQSISSPLDNKLNVEDVDSTKNRKLIDDY 120
Qy 121 DSTKSLDGHKFDQDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180
Db 121 DSTKSLDGHKFDQDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPKTSWTENQAGIKPKVTPMAIQDGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPKTSWTENQAGIKPKVTPMAIQDGL 240
Qy 241 AKGENDETYSNTLTITNGLERTKYSEDNFELQVFNFNFYALLKSIDSEKEKEKTLI 300
Db 241 AKGENDETYSNTLTITNGLERTKYSEDNFELQVFNFNFYALLKSIDSEKEKEKTLI 300
Qy 301 TIMKTLIDFVQMKVYGTISPEGVSYLENLDEMIALQTKLNKLNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVQMKVYGTISPEGVSYLENLDEMIALQTKLNKLNATDNISKLPAPSE 360
Qy 361 KSHHEETDSTKEAAKMEYGLKSDTKDDNSNPGKTDPEKGTAEVLAIRKNI EWLK 420
Db 361 KSHHEETDSTKEAAKMEYGLKSDTKDDNSNPGKTDPEKGTAEVLAIRKNI EWLK 420
Qy 421 KHDKGNKEDYDLKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468

Db 421 KHDKGNKEDYDLKMRDFINKQADAYVEKGLDKEEAIAIKRIYSSL 468
RESULT 14
AAO13905
ID AAO13905 standard; Protein; 567 AA.
XX AAO13905;
AC AAO13905;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polypeptide SEQ ID NO 27797.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation.
XX
OS Homo sapiens.
XX
XX WO200164835-A2.
XX
PD 07-SEP-2001.
XX
XX 26-FEB-2001; 2001WO-US04927.
XX
XX 28-FEB-2000; 2000US-0515126.
XX
XX 18-MAY-2000; 2000US-0577409.
XX
XX (HYSE-) HYSEQ INC.
XX
XX Tang YT, Liu C, Drmanac RT;
XX
XX WPI; 2001-514838/56.
XX
XX N-PSDB; AAI93836.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing
XX diagnosing and treating e.g. leukaemia, inflammation and immune
XX disorders -
XX
XX Claim 20; SEQ ID NO 27797; 1399pp + Sequence Listing; English.
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX cytokine, cell proliferation or cell differentiation or which may induce
XX production of other cytokines in other cell populations. The
XX polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX peptide therapy. The polypeptides have various cytokine-like activities,
XX e.g. stem cell growth factor activity, haematopoiesis regulating
XX activity, tissue growth factor activity, immunomodulatory activity and
XX activin/inhibin activity and may be useful in the diagnosis and/or
XX treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX inflammation.
XX Note: The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 567 AA;

Query Match 100.0%; Score 2399; DB 22; Length 567;
Best Local Similarity 100.0%; Pred. No. 5.3e-146;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTTWLVLPQAPFKPGGQDLSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 60
Db 100 MGFLGTGTTWLVLPQAPFKPGGQDLSLHNRSLAERPLNEQIAEAEEDKIKKTYPP 159
Qy 61 ENKPGQSNYSFVNDLNLKAI TEKEKERQSISSPLDNKLNVEDVDSTKNRKLIDDY 120
Db 160 ENKPGQSNYSFVNDLNLKAI TEKEKERQSISSPLDNKLNVEDVDSTKNRKLIDDY 219
Qy 121 DSTKSLDGHKFDQDPPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGL 180

Db 220 DSTKSGLDHKKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 279
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVPMPMAAIOQGL 240
Db 280 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVPMPMAAIOQGL 339
Qy 241 AKGENDETIVSNTLTLTNGLERRTKTYSEDNFEELQYFPNFVALLKSIDSEKEAKEKETLI 300
Db 340 AKGENDETIVSNTLTLTNGLERRTKTYSEDNFEELQYFPNFVALLKSIDSEKEAKEKETLI 399
Qy 301 TIMKTLLIDFVKMVKYGTISPEEGVSYLENLDEMIALQTKNLEKATDNTSKLFPAPSE 360
Db 400 TIMKTLLIDFVKMVKYGTISPEEGVSYLENLDEMIALQTKNLEKATDNTSKLFPAPSE 459
Qy 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 420
Db 460 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 519
Qy 421 KHDKKGKNEKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
Db 520 KHDKKGKNEKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 567

RESULT 15

AAW93910
ID AAW93910 standard; Protein; 468 AA.

XX AC AAW93910;
XX DT 06-NOV-2001 (first entry)
XX DE Human polypeptide, SEQ ID NO: 4061.
XX KW Human; full length cDNA; cDNA synthesis; oligo-capping.
XX OS Homo sapiens.
XX PN EP1130094-A2.
XX PD 05-SEP-2001.
XX PF 07-JUL-2000; 2000EP-0114089.
XX PR 08-JUL-1999; 99JP-0194486.
XX PR 11-JAN-2000; 2000JP-0118774.
XX PR 02-MAY-2000; 2000JP-0183765.
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Nishikawa T, Isogai T, Hayashi K, Ishii S, Kawai Y;
XX PI Wakamatsu A, Sugiyama T, Nagai K, Kojima S, Otauki T, Koga H;
XX DR WPI; 2001-524255/58.
XX DR N-PSDB; AAK94872.

830 Primers useful for synthesizing full length cDNA clones and their use in genetic manipulation

PS Claim 8; SEQ ID NO 4061; 1380pp + sequence listing; English.

XX The invention relates to primers for synthesising full length cDNA clones. 830 cDNA molecules encoding a human protein have been isolated and nucleotide sequences of 5'- and 3'-ends of the cDNA molecules have been determined. Primers for synthesising the full length cDNA are useful for clarifying the function of the protein encoded by the cDNA. The full length clones were obtained by construction of full length enriched cDNA libraries that were synthesised by the oligo-capping method. The primers enable the production of the full length cDNA easily without any special methods. The present sequence is a polypeptide encoded by a full length human cDNA of the invention.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in CD-ROM format directly from EPO.

XX SQ Sequence 468 AA;
Query Match 99.7%; Score 2392; DB 22; Length 468;
Best Local Similarity 99.8%; Pred. No. 1.2e-145;
Matches 467; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 MGFLGTCTWLVLPLOAFPKPGGSDKSLHNRRLSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTCTWLVLPLOAFPKPGGSDKSLHNRRLSAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVNDNLNLLKAITKEKTEKERQSRSSPDLNKLNVEDVDSTKNRKLIDDY 120
Db 61 ENKPGQSNYSFVNDNLNLLKAITKEKTEKERQSRSSPDLNKLNVEDVDSTKNRKLIDDY 120
Qy 121 DSTKSGLDHKKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHKKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVPMPMAAIOQGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVPMPMAAIOQGL 240
Qy 241 AKGENDETIVSNTLTLTNGLERRTKTYSEDNFEELQYFPNFVALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETIVSNTLTLTNGLERRTKTYSEDNFEELQYFPNFVALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLLIDFVKMVKYGTISPEEGVSYLENLDEMIALQTKNLEKATDNTSKLFPAPSE 360
Db 301 TIMKTLLIDFVKMVKYGTISPEEGVSYLENLDEMIALQTKNLEKATDNTSKLFPAPSE 360
Qy 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEEAAKMEKEYGSLKSDTKDDNSNPGKTDPEPKGTEAYLEAIRKNIEWLK 420
Qy 421 KHDKKGKNEKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
Db 421 KHDKKGKNEKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468

Search completed: January 2, 2004, 11:01:00
Job time : 48 secs

THIS PAGE BLANK (033710)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 10:59:32 ; Search time 21 Seconds
(without alignments)
942.928 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGTLVLVLPIQAF.....EKGILDKERAEAKRIYSSL 468

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA.*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pap.*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pap.*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pap.*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pap.*
5: /cgn2_6/ptodata/1/iaa/PTCUS_COMB.pap.*
6: /cgn2_6/ptodata/1/iaa/backfiles.pap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	179	7.5	1786	3	US-08-973-462-8
2	172.5	7.2	1164	3	US-08-923-992A-2
3	169	7.0	1098	3	US-08-923-992A-8
4	168	7.0	1164	3	US-08-923-992A-10
5	167.5	7.0	957	4	US-09-914-259-16
6	166	6.9	1104	3	US-08-923-992A-4
7	166	6.9	1128	3	US-08-923-992A-6
8	166	6.9	2662	4	US-09-595-684B-31
9	163.5	6.8	956	4	US-09-914-259-17
10	162	6.8	8991	4	US-08-714-741-32
11	159.5	6.6	816	2	US-08-533-306A-6
12	159.5	6.6	816	2	US-08-742-923A-6
13	158.5	6.6	885	2	US-08-533-306A-4
14	158.5	6.6	885	2	US-08-742-923A-4
15	156.5	6.5	1087	4	US-09-914-259-12
16	156	6.5	1354	3	US-08-685-871-2
17	154.5	6.4	1231	4	US-08-714-741-41
18	153.5	6.4	984	1	US-08-242-932-2
19	153.5	6.4	984	1	US-08-714-481-2
20	153.5	6.4	984	5	PCT-US95-06111-2
21	153	6.4	1588	5	PCT-US93-07261-11
22	153	6.4	1663	5	PCT-US93-07261-16
23	153	6.4	2285	4	US-09-308-375-2
24	150.5	6.3	1010	4	US-09-134-001C-5178
25	147	6.1	534	4	US-09-103-664A-2
26	146.5	6.1	1388	2	US-08-685-576-1
27	146.5	6.1	1388	2	US-08-685-576-4

28	145.5	6.1	652	4	US-08-559-896B-2
29	145	6.0	588	4	US-08-714-741-42
30	144.5	6.0	343	3	US-08-937-271-17
31	144	6.0	10182	4	US-09-134-001C-3159
32	143.5	6.0	864	4	US-08-714-741-40
33	143	6.0	3878	4	US-09-914-259-11
34	142	5.9	967	4	US-09-914-259-21
35	142	5.9	1032	4	US-09-914-259-26
36	141.5	5.9	1146	4	US-08-914-999-6
37	141	5.9	1507	3	US-08-929-329-5
38	139.5	5.8	963	4	US-09-914-259-20
39	139	5.8	1939	4	US-09-310-187A-1
40	139	5.8	3111	2	US-08-460-309-4
41	139	5.8	3111	2	US-08-125-077-4
42	138	5.8	630	3	US-08-973-462-9
43	138	5.8	1177	4	US-09-134-001C-5106
44	138	5.8	1234	4	US-09-592-054-8
45	137.5	5.7	586	2	US-08-630-822A-70

ALIGNMENTS

RESULT 1
US-08-973-462-8
; Sequence 8, Application US/08973462B
; Patent No. 6191270
; GENERAL INFORMATION:
; APPLICANT: DRUILHE, PIERRE
; APPLICANT: DAUBERSIES, PIERRE
; FILE OF INVENTION: MALARIAL PRE-ERYTHROCYTIC STAGE POLYPEPTIDE MOLECULES
; CURRENT FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: US/08/973.462B
; EARLIER FILING DATE: 1998-02-06
; EARLIER APPLICATION NUMBER: PCT/FR96/00894
; EARLIER FILING DATE: 1996-06-12
; EARLIER APPLICATION NUMBER: FR 95/07007
; EARLIER FILING DATE: 1995-06-13
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8
; LENGTH: 1786
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Polypeptide
US-08-973-462-8

Query Match	7.5%; Score 179; DB 3; Length 1786;
Best Local Similarity	21.5%; Pred. No. 0.00012;
Matches	118; Conservative 96; Mismatches 197; Indels 138; Gaps 25;
Qy	27 QDKSLHNRLEISAE--RPLNEQIAEAEEDKIKYTPPENKPGQSN-----YSFVDNL-- 75
Db	957 EENAVESNENVAENLEKLNFTVTLVDKVEETVEISGESLENNEMDKAFSEIFDNVKG 1016
Qy	76 ---NLK-----AITEKEKIEKEROISRSPLDNKLN-----VEDVDSTK-- 112
Db	1017 IQENLLTGMRFSIETSIQSEKVDL--NENVSSILDNIENMKGLLNKLENISSTEGV 1075
Qy	113 -----NRKLDDYD-----STKGLDHKFDQDDPDLHQLDGTPLTAEDI- 151
Db	1076 QETVTEHVEQNYYVDVDPAMKQDFLGILNEAGLKEMFNFLED--VPKSESDVITVEIK 1134
Qy	152 -----VHKIARIYEENDRAVDFKVLKLNGLLTESQAHTLEDEVAEVLQKLSKE 204
Db	1135 DEPVQKEVEKETVSIIEEMEENIVD-----VLEEERDLTKMDAVEESIEIS 1183
Qy	205 ANNYEDPNKPTSTWENQAGKIPKVTPTMAAIQDLAKGENDETIVSNLTLTINGLERRTK 264
Db	1184 SDSKEE-----TES-----IKDKKDVSVLVVEVDNDMDSEVKVLELKNMBELMK 1231
Qy	265 TYSEDN-----FEELQYFPNFYA--LLKSIDSEKAK-----EKETLITIMK 304

```

Db 1232 DAVEINDITSLIBETQELNVEADLIKDMKLEKLEKALSDSKSEIIDAADDTLEKVE 1291
QY 305 -----TLIDFVKMVKYGTISPE--EGVSYLENLDEMIALQTK--NKLKRNATDNISKL 354
Db 1292 EHDITITLDEV---VELKDVDEKIEKVSOLKLEEDILKEVKEIKELSEILEDYKEL 1348
QY 355 FPAPSEKSHETDSTKEEAAKMEYKSLKSDTKDDNSNPGGKTDEPKGKTEAYLEAIRK 414
Db 1349 KTIETDILEEKEIKERKOHFEKFEFEAEIKOLEADILKEVSSLEVEBEKKLEEVHE-LKE 1407
QY 415 NIEWLKHKDK--KNGKEDYLSKWRDFNKQADAY---VEKGILDK-----EEA 458
Db 1408 EVEHIIISGDAHIKGLEED-DLEEVDDLKSGSLDMLKGMELGDMDKESLEDVTTKLGSRV 1466
QY 459 EAIKRIYSS 467
Db 1467 ESLKDVLS 1475

RESULT 2
US-08-923-992A-2
; Sequence 2, Application US/089233992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-2

Query Match 7.28; Score 172.5; DB 3; Length 1164;
Best Local Similarity 21.68; Pred. No. 0.0002;
Matches 124; Conservative 77; Mismatches 189; Indels 185; Gaps 25;

QY 36 LSAREPLNE-----QIAEABEDIKIKTYPPENKPGQSNYSVDNLLNLKAIT-EKEK 86
Db 149 LELENQFNETNRLHLIKHQHEVEKDKAK-----QOKTLKQSDTKVDLSNIDKELNHQKSO 204
QY 87 IEK-ERQSTRSSPLDNKL-----NVEDVDSTKRNKLIDYDSTKSGLDHK 130

```

```

Db 205 VEKMAEQKGITNEKDSMLKKIEDIRKQAQAKKDEAEVKVREELKGLFSSTKAGLDQE 264
QY 131 FODDPDGLHQLDGTPLTAEDIVHKI-----AARIYEENDRAVPDKIVSKLLNLG 179
Db 265 IQE-----HVKKET--SSEENTQKVDEHYANSLQNLAKSLEELDKATTNEQATQVKNQF 317
QY 180 L-----ITESQAHTLEDEVAEV-----LOKLIKE--ANNYEE 210
Db 318 LENAQKLKEIOPLIKETNVKLYKAMSESLQVEKELKHNSANLEDLVAKSKEIVREYEG 377
QY 211 DPNKPTSWTE-----NOAGKIPEKVTPTMAAIOQDGLAKGENDETWSN 251
Db 378 KLNQSKNLPQLKEEAHSLKQVVEDFRKFKTSEQVTPKRVKRDLAANNQ---Q 434
QY 252 TLITLNGLERTKTYSNEDNFEELQYFPNFYALLXSIDSEKAKEKETLITINKTILDFVK 311
Db 435 KIETVSPENITVYEGED-----VKFTVTAKS-DS-----KTLDFSD 471
QY 312 MMVKG-TISPPEGVSVLENLD-----EMIALQTK-----NKLEKNATD 349
Db 472 LTKYNSVSDRISTNYKNTDHNHIAITIKNLKLNESQVTLKAKDDSGNVVETFTI 531
QY 350 NISKLPAPSEKSHETDSTKEEAA-----KMEKYGSLKD 385
Db 532 TVQKKEEKQVPTPEQKDSKTEKVPQEPKNDKNQLQELIKSAQOQLEKLEKAIKELME 591
QY 386 STKDDNSNPGGKTDEPKGKTEAYLEAIRKNTLEWLK-----HDKKNGKEDYLSKWRDFN 441
Db 592 Q-PEIPSNP--EYGIQKSIWESQKEPQEAITSFKIIGDSSSKYYTEHYFNKYKSDPMN 648
QY 442 KQADAYVEKGILDKKEAE-----AIKRIYSS 467
Db 649 YQLHAQME--MLTRKVVQVMKYPDNAEIKKIFES 681

RESULT 3
US-08-923-992A-8
; Sequence 8, Application US/089233992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 8:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-2

```

```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 1098 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-8

Query Match          7.0%; Score 169; DB 3; Length 1098;
Best Local Similarity 19.3%; Pred. No. 0.00034;
Matches 104; Conservative 80; Mismatches 182; Indels 174; Gaps 19;

QY 25 GSQKSLHNLRLSARPLNEQIAEAEEDKIKKTYPPNKPQGSNSYF---VDNLNLLKAI 81
Db 74 GKREKQLQWKNNLKNVDNLTLSHEQKNBEFKTKIDETNDSALLELENGFNENRLLHI 133
QY 82 TEKEKIEROSIRSSPL---DNKLNVEDVDSTKNR-----KLIDQ--- 119
Db 134 KQHEBEVKDKAKAQKQTKLKQSDTKVDLSNIDKELNHQKSQAGITNEDKSMLKIEDIR 193
QY 120 -----YDSTKSGLDHKFQDDPGDLHQLDGTPLTAED 155
Db 194 KQAQPKKEDAEVKVREELGKLFSSKAGLDQEIQE-----HVKKET--SSEENTQKVD 246
QY 156 -----AARIYEENDRAVDFKIVSKLNLGL-----ITESQA 186
Db 247 EHYANSLQNLQAQKSLLELDKATTEQATQVKNQFLENAQKLKEIQPLIKETNVKLYKAMS 306
QY 187 HTLEDEVAEV-----LQKLISKE---ANNVEEDPNKPTSWTE----- 220
Db 307 ESLEQVEKELHNSANLELDVAKSKEIVREYEGKLNQSKNLPKQLBEBAHSLKQKV 366
QY 221 ---NOAGKIPKVTMPMAAIOQLAKGENDETVSNLTITNGLERRTKTSYSDNFEELOYF 277
Db 367 EDFRKKPTSQVTPKPKRLDLAANNQ---QKIELTVSPENITVYEGD----- 415
QY 278 PNFVALLKSISEKEAKEKETLITIMKTLIDFPVGMVKYG-TISPBEVSVYLENLD 332
Db 416 VKFTVTKAS-DS-----KTLDFSDLLTKYNFSDRISTNTYKNTDNHKI 460
QY 333 -----EMIALQTK---NKLEKATDNISKLPFPAPSEKSHSEETDSTKEAAK 375
Db 461 AEITIKNLKNESQTVTLKADSGNVVEKFTTITVQKKEKQVPTPEQKDSKTEKVP 520
QY 376 MEKEVGLSKDSTKDNNSPGGKTEPKGTETAYLEAIRKNIEWLWKHDKGNKEDYDLISK 435
Db 521 QEPK-----SNDKNQLQELIKSAQAELEKLEKAIKELMEQPEIFSNPEYGIQK 568

RESULT 4
US-08-923-992A-10
; Sequence 10, Application US/08923992A
; Patent No. 6280738
; GENERAL INFORMATION:
; APPLICANT: Tai, Joseph Y.
; APPLICANT: Blake, Milan S.
; TITLE OF INVENTION: No. 6280738-Iga Fc Binding Forms of the Group B
; TITLE OF INVENTION: Streptococcal Beta Antigens
; NUMBER OF SEQUENCES: 34
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
; STREET: 1100 New York Avenue, N.W., Suite 600
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/923,992A
; FILING DATE: 05-SEP-1997

```

```

; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/024,707
; FILING DATE: 06-SEP-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Esmond, Robert W.
; REGISTRATION NUMBER: 32,893
; REFERENCE/POCKET NUMBER: 1438.0140001/RWE
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 371-2600
; TELEFAX: (202) 371-2540
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1164 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-923-992A-10

Query Match          7.0%; Score 168; DB 3; Length 1164;
Best Local Similarity 19.0%; Pred. No. 0.00044;
Matches 121; Conservative 85; Mismatches 177; Indels 254; Gaps 25;

QY 25 GSQKSLHNLRLSARPLNEQIAEAEEDKIKKTYPPNKPQGSNSYF---VDNLNLLKAI 81
Db 105 GKREKQLQWKNNLKNVDNLTLSHEQKNBEFKTKIDETNDSALLELENGFNENRLLHI 164
QY 82 TEKEKIEROSIRSSPL---DNKLNVEDVDSTKNR-----K 115
Db 165 KQHEBEVKDKAKAQKQTKLKQSDTKVDLSNIDKELNHQKSQVETMAEQIGITNEDKDSMLK 224
QY 116 LIDQ-----YDSTKSGLDHKFQDDPGDLHQLDGTPLTAED 150
Db 225 KIEDIRKQAQADKEDAEVKVREELGKLPSTKAGLDQEIQE-----HVKKET--TSEE 277
QY 151 IVHKI-----AARIYEENDRAVDFKIVSKLNL-----LGLITE----- 183
Db 278 NTQKVDHYPNSLQNLQAQKSLLELDKATTEQATQVKNQFLENAQKLKEIQPLIKETNVK 337
QY 184 -----SQAHTLEDEVAEVFLKLSKE---ANNVEEDPNKPTSWTE--- 220
Db 338 LYKAMSSLEQVEKQLKHNQAN--LEDLVAK-----SKEIVREYEGKLNQSKNLPKQ 390
QY 221 -----NOAGKIPKVTMPMAAIOQLAKGEND-----TVS--NTLTITNG 258
Db 391 LEEBAHSLKQVVEDFRKKPTSQVTPKPKRDLAANNQKIELTVSPENITVYEG 450
QY 259 LERRTKTSYSDNFEELOYFPNFYALLKSISEKEAKEKETLITIMKTLIDFPVGMVKYG- 317
Db 451 -----EDLK-----FTLTAKSDS-----KTLDFSDLLTKYNP 478
QY 318 TISPBEVSVYLENLD-----EMIALQTK---NKLEKATDNISKLP 356
Db 479 SVSDRISTNTYKNTDNHKAIBITIKNLKLNESQTVTLKADSGNVVQKTFITVQKKEE 538
QY 357 APSEKSHSEETDSTKEAAKMEKEVGLSKDSTKDNNSPGGKTEPKGTETAYLEAIRKNI 416
Db 539 KQVPTPEQKDSKTEKVPQEPK-----SNDKNQLQELIKSAQAELEKLEKAI 586
QY 417 EWLKXKDKGKNKEDYDLISK-----AIKRIYSS 467
Db 587 KELMEQPEIFSNPEYGIQKSIWESQKEPIQEAITSFKKIIGDSSSKYYTTHYFNKYKSD 646
QY 440 INKQADAYVEKGLDKEEAE-----AIKRIYSS 467
Db 647 MNYQLHAQME--MLTRKVVQYINKYPONABIKKIFES 681

RESULT 5
US-09-914-259-16
; Sequence 16, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:

```


US-08-923-992A-6
 ; Sequence 6, Application US/08923992A
 ; Patent No. 6280738
 ; GENERAL INFORMATION:
 ; APPLICANT: Tai, Joseph Y.
 ; APPLICANT: Blake, Milan S.
 ; TITLE OF INVENTION: No. 6280738-IgA Fc Binding Forms of the Group B
 ; TITLE OF INVENTION: Streptococcal Beta Antigens
 ; NUMBER OF SEQUENCES: 34
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Sterne, Kessler, Goldstein & Fox P.L.L.C.
 ; STREET: 1100 New York Avenue, N.W., Suite 600
 ; CITY: Washington
 ; STATE: D.C.
 ; COUNTRY: USA
 ; ZIP: 20005
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/923,992A
 ; FILING DATE: 05-SEP-1997
 ; CLASSIFICATION: 536
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US 60/024,707
 ; FILING DATE: 06-SEP-1996
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Esmond, Robert W.
 ; REGISTRATION NUMBER: 32,893
 ; REFERENCE/DOCKET NUMBER: 1438.0140001/RWE
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (202) 371-2600
 ; TELEFAX: (202) 371-2540
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1128 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-08-923-992A-6

Query Match 6.9%; Score 166; DB 3; Length 1128;
 Best Local Similarity 19.3%; Pred. No. 0.00059;
 Matches 119; Conservative 86; Mismatches 200; Indels 210; Gaps 23;

Qy 25 GSQKSLHNRLSARPLNEQIAEAEEDKIKKTPPENKPGQSNYSP---VDNLNLKAI 81
 Db 69 GKREKQLOQWKNLNDVNTILSHEQKNEFKTKIDETNDSALLELENQFNETNRLHI 128
 Qy 82 TEKEKIERQSISSPL---DNKLNVEDVDSTKNR-----K 115
 Db 129 KQHEVEKDKAKQKQTLKQSDTKVLSNIDKELNHQKSPVEKMAEPKGIINEDKDSMLK 188
 Qy 116 LIDD-----YDSTKSLDHPKQDDPDGHLQDGLTPLTAED 150
 Db 189 KIEDIRKQAQADKEDAEVKVREELKGLFSGTRAGLDQE-----IHEVHKETSSEE 241
 Qy 151 IVHKI-----AARIYEENDRAVDFKIVSKLLNLGL----- 180
 Db 242 NTQKVDEHYANSLONLAQKSLLEEDKATTNQCATQVKNQFLENACKLQKEMQPLIKETNVK 301
 Qy 181 ITESQAHTLEDEAVR-----LQKLISKE---ANNYEEDPNKPTSWE----- 220
 Db 302 LYKAMESLEQVEKELKNSANLEEDLVAKSKEIVREYEGKLNQSKNLPKQLQLEEAHS 361
 Qy 221 -----NOAGKIPKVTPTMAAIQDGLAKGENDETVSNLTTLTNGLERRTKTSYDNF 271
 Db 362 KLKQWVEDFRKKFTSEQVTPKRVKRDLANENQ---QKIELTVSPENITVVEGED-- 416
 Qy 272 EELQVFPNFPYALLKSIDSEKAKEKETLITMTKLIDFVKRMVKY-TISPEGVSYLEN 330

Db 417 -----VKFTVTAKS-DS-----KTTLDPSDLLTKYNPSVSDRISTNYKTN 455
 Qy 331 LD-----EMIALQTK-----NKLEKNATDNISKLPAPSEKSHETDST 369
 Db 456 TDNHKIAEITIKNLKLNESQTVTLKAKDDSGNVVVEKFTTITVQKKEKQVPKTEQKDSK 515
 Qy 370 KEAAA-----KMEKEYGSLKSDTKDDNSNPGKTDPEPKGT 405
 Db 516 TEEKVQPEPKSNDKNQLOELIKSAQOELEKLEKAIKELMEQ-PEIPSNP--EYGIQKSIW 572
 Qy 406 EAYLEATRKNIEWLK-----HDKKGNKEDYDLSQMRDFINKQADAVYVEKGLDKEEAE-- 459
 Db 573 ESQKEPIQEAITSFKKLIIGDSSSKYYTEHYFNKYKSDPMYQLHAQME--MLTRKVVQYM 630
 Qy 460 -----AIKRIYSS 467
 Db 631 NKYPDNAEIKKIPES 645

RESULT 8
 US-09-595-684B-31
 ; Sequence 31, Application US/09595684B
 ; Patent No. 6544766
 ; GENERAL INFORMATION:
 ; APPLICANT: Beraud, Christophe
 ; APPLICANT: Ohashi, Cara
 ; APPLICANT: Sakowicz, Roman
 ; APPLICANT: Vaisberg, Eugeni
 ; APPLICANT: Wood, Kenneth
 ; APPLICANT: Yu, Ming
 ; TITLE OF INVENTION: Human kinesins and methods of producing
 ; TITLE OF INVENTION: and purifying human kinesins
 ; FILE REFERENCE: cytop036
 ; CURRENT APPLICATION NUMBER: US/09/595,684B
 ; PRIOR FILING DATE: 2002-06-24
 ; PRIOR APPLICATION NUMBER: 09/295,612
 ; PRIOR FILING DATE: 2000-04-20
 ; NUMBER OF SEQ ID NOS: 105
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 31
 ; LENGTH: 2662
 ; TYPE: PRT
 ; ORGANISM: Human
 ; US-09-595-684B-31

Query Match 6.9%; Score 166; DB 4; Length 2662;
 Best Local Similarity 21.8%; Pred. No. 0.0019;
 Matches 119; Conservative 103; Mismatches 175; Indels 150; Gaps 28;

Qy 26 SQKSLHNRLSARPLN---EQIAEAEEDKIKKTPPENKPGQSNYSFVDNLNLKAIT 82
 Db 1101 AQEK---NHAIKKEGELSRTCRLAEVEEKLKESQQLQEKQQL-----LNVQEEMS 1150
 Qy 83 EKKEKIERQSISSPLDNKLNVEDVDSTK---NRKLIIDDYDSTKS-----GLDHK 130
 Db 1151 EMOKKINEIENKLNKELTLEHMETERLELAQKLNENYEVKSTKERRKVLKELQKS 1210
 Qy 131 FODDPDGL-----HOLDGTPL-TAEDI---VH-KIAARIYEENDRAVDFKIVSKLLNLGL 180
 Db 1211 PETERDHLRVIYRIEATGLQTEELKIAHILKHEOETIDELRRSVSEK-TAQIINTQD 1269
 Qy 181 ITESQAHTLEDEAVR---QKLIS--KEANNYEEDPNKPTSMTENQAGKIPKVTPTMAA 235
 Db 1270 LEKS---HTKLQEEIPLVHBEQQLPNVVKYSETCETNLELLTEQSTTK---DSTTLAR 1324
 Qy 236 IQ-----DGLAK-GENDETVENTL-----TLTNGLERRTK 264
 Db 1325 IEMERLRLNEKFOESQBEIKSLTKERDNLTKIKEALEVQDLKHEHIRETLAKIQESQSK 1384
 Qy 265 TYSEDNFEEL-----QYFPNFPYALLK-----SIDSEK-AKEKE 297
 Db 1385 QEOSLNKKEKDNETTKIVSEMEQFKPKDSALLRIEIMLGLSKRLQESHDEMKSVAKEKD 1444

QY 238 TLITIMKTLDIVKMMVYGTISPEGVSYLENLDMIA--LQTKNLE-----KNATD 349
Db 1445 DLQRLEVL-----QSEDLKENIYKIVAKHLETEELKVAHCCLEQEE 1490
QY 350 NISKLFPAPSEKSHETSTKEAAKMEKYGSLKSDKSTDDNSNPGKTDPEKGTKE--- 406
Db 1491 TINELRVNLSEK-----ETEISTIQKLEAINDKLQ-----KIQEIYEKEQLN 1535
QY 407 -AYLEAIRKNLEWKHKDKGNKEDYD-----SKRDPINKQADAYVEKGLDKKEAEAI 461
Db 1536 IKQISEVENNELKQFKEHRKAKDSALQSTESKMLELTNRLQESQEEIQIMIKEK-BEM 1594
QY 462 KRIYSSL 468
Db 1595 KRVQEAL 1601

RESULT 9
US-09-914-259-17
; Sequence 17, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 17
; LENGTH: 956
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-914-259-17

Query Match 6.8%; Score 163.5; DB 4; Length 956;
Best Local Similarity 20.7%; Pred. No. 0.00072;
Matches 106; Conservative 91; Mismatches 169; Indels 147; Gaps 25;

QY 25 GSQDKSLHNR-----ELSAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSFVDNLLNLL- 78
Db 321 GORAKTIKNTVSNVLELTAE-----EWKKYKEKEKNAKSLVLOHLEWEL 367
QY 79 -----KAITEKIE-KERQSISSPLDNKLVED-----VDSTKNKLDIDYDSTKSL 127
Db 368 NRWNGEAVPEDEQISAKDHKSL--EPCDTPIIDNITPVVDGISAEB--EKYDEEITSL 423
QY 128 DHKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEBND-----RAVFDKIVSKLINLGLIT 182
Db 424 YRQLDDKDEINO-----QSQAELKQLQMLDQDELLASTRADYDKIQBELTRLQIEN 476
QY 183 ESQAHTLEDEVAELVLOKLSKEANNYE--EDNKPITSWTENQAGKIPEKTPMAAIQDG 239
Db 477 EA-----AKDEVEKVLQAL-BELAVNYDQKSQVEDKTRANEQLTDLAQKTTTLTTTQRE 531
QY 240 LAK-----GENDETVSNTLTNLGLERR-----TKT 265
Db 532 LSQQLSNHKKRATETINLLKDLGEIGITGND-----VKTLADVNGVIEEFTWARL 588
QY 266 YSDENFEELYFNFVALLKS--IDSEKAKEKETLITIMKTLDIVKMMVYGTISPEE 323
Db 589 YISKMSEVKSLVNRSKQLESQMSDNKRKNWASERELAAQCLLISQHEAKISLT----- 643
QY 324 GVSYLENLDMIALOTNKNLEK--NATDNISKLPAPSEKSHET--DSTKEBAKMEK 378
Db 644 --DYQNMW-----QKRRQLESQDSLSEELAKL--RAQEKMEHVSFQDKKEKHLRLQ- 693
QY 379 EYGLSKDSTKDDNSNPGKTDPEKGTKEAYLEAIRKNLEWKHKDKGNKEDYDLSKORD 438
Db 694 -----DAEEVYKALEQQWESHR-----EAHQKQ-----LSRLRD 722

QY 439 FINKQADAYVEKGILDKE---FAEAKRIYSSSL 468
Db 723 EIEKQRIIDEIRDLNQKLEQEERLSSDYNNKL 755

RESULT 10
US-08-714-741-32
; Sequence 32, Application US/08714741
; Patent No. 6500613
; GENERAL INFORMATION:
; APPLICANT: Briles, David E.
; APPLICANT: McDaniel, Larry S.
; APPLICANT: Swiatlo, Edwin
; APPLICANT: Yother, Janet
; APPLICANT: Crain, Marilyn J.
; APPLICANT: Hollingshead, Susan
; APPLICANT: Tart, Rebecca
; APPLICANT: Brooks-Walter, Alexis
; TITLE OF INVENTION: PNEUMOCOCCAL GENES, PORTIONS THEREOF,
; TITLE OF INVENTION: EXPRESSION PRODUCTS THEREFROM, AND USES OF SUCH GENES,
; TITLE OF INVENTION: PORTIONS AND PRODUCTS
; NUMBER OF SEQUENCES: 47
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Curtis, Morris & Safford, P.C.
; STREET: 530 Fifth Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/714,741
; FILING DATE: 16-SEP-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Frommer Esq., William S.
; REGISTRATION NUMBER: 25,506
; REFERENCE/DOCKET NUMBER: 454312-2460
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 840-3333
; TELEFAX: (212) 840-0712
; INFORMATION FOR SEQ ID NO: 32:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 8991 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: amino acid
US-08-714-741-32

Query Match 6.8%; Score 162; DB 4; Length 8991;
Best Local Similarity 18.2%; Pred. No. 0.02;
Matches 88; Conservative 110; Mismatches 184; Indels 102; Gaps 19;

QY 26 SQDKSLHNRSLAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSFVDNLLNLLKAITEKE 85
Db 2024 AQEKADYQRIIEKAKEQASLEQOEANKYQLKLYLDGRLNSVLLKKEEAE 2083
QY 86 KIKERQSISSPLDNKLV-----NVEDVDSTKNKLDIDYDSTKSLDGHKFD----- 133
Db 2084 KDKKEQA-----EFNKIRREIVVPNPQLEEMARKSEV--VKAKESGLVKRVEEAEKV 2136
QY 134 -----DPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVS-KLINLGLITESQ 185
Db 2137 TEARQKLAERAKEVVLQPTRVENEVHKLQKQNK--KMLTSLASVALGAGLVTSP 2194
QY 186 AHTLEDEVAELVLOK-----LISKEANN-----YBEDPNKPTSWT 219
Db 2195 TFVRAEESQWVEKSEKKEKYEAKAKADTAKKDYETAKKAEADQKKYBDQKR-----T 2250

Qy	220	ENOAGKIPE---KVTPMA-----AIQGLAKGENDETVSNTLT-LTNGLERRTK	264
Dd	2251	EEKARKEEASQKLIDVALVQNAYREYEVQNRSKYSKDADYQKQLTEVDSKIEARK	2310
Qy	265	TYS--EDNFELOYFPNFYALL---KSIDSEKEAKEKETLITIMTKLIDFVKMMVKYT	318
Dd	2311	EQOQLQNNFNVR-----AVPADPTCVGXDXRMWNKKMILTSLSV-----AILGAGX	2359
Qy	319	ISPEEGSVYLENDIMATALOTKNLEKANTDNISKLPAPSEKSHSHEETSTKEAAKWEK	378
Dd	2360	VTSQPTXVRAEEAPQVV---EKSLERKYBEEAKY-----DAAKKDYDEAKKAABEQK	2411
Qy	379	EYGS�DKSTDSDNSPGCKTDEPKGTAEVLAIRKNI EWLKKHDDKGN---KEDVDLSK	435
Dd	2412	KYEDQKKT EEKAEKAKAASEEIAKATEEVQKAVLDYITAIRHNDSKGTSAAEAENKAK	2471
Qy	436	MRDF 439	
Dd	2472	ERDY 2475	

```

RESULT 11
US-08-533-306A-6
; Sequence 6, Application US/08533306A
; Patent No. 5837457
; GENERAL INFORMATION:
; APPLICANT: Liu, Pu
; APPLICANT: Collins, Francis S.
; APPLICANT: Siciliano, Michael J.
; APPLICANT: Claxton, David
; TITLE OF INVENTION: Markers for Detection of Chromosome 16
; TITLE OF INVENTION: Rearrangements
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Harness, Dickey & Pierce, P.L.C.
; STREET: P.O. Box 828
; CITY: Bloomfield Hills
; STATE: MI
; COUNTRY: USA
; ZIP: 48303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/533.306A
; FILING DATE: September 25, 1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, Deann F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869COB
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 816 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-533-306A-6

```

Qy	84	KKIKERQSIR-----SSPLDNKNVDDVDSTKN-----RKLIIDDYDSTKSGLDH---KF 133
Dd	216	LRLQLEERNLSIQDQLDBEMEAQNLERHIISTLNISLSDSKKKLQDFASTVEALEEGKKRF 275
Qy	132	QQDPDGHLHOLDGTPLTAEIDIVHKTAARIYEENDRAVFD-----KITVKLLNIGLITESQA 186
Dd	276	QKETENUTQQYEEKAAAYDKLEKTKNRLLQQELDLVVLDNDNQRLVSNL-----EKKQ 328
Qy	187	HTEDEVAEVQLKLISKANNYBEDPNKPSTSWTNQAGKIPEKVTFPMA-AIODGLAKGEN 245
Dd	329	RKPDQLAEE-EKNIS---SKYADEROR---AEAAREKETKALSARALBEALEAKEE 379
Qy	246	DETUSNTILTLNGUERTKTVSED-----NFPELOYFPNPFYALLKSIDSKEKAKEKETLIT 301
Dd	380	LERTNMKLK---AEMEDLVSSKDQVGKNVHEL-----EKSKRALETOME 420
Qy	302	IMKTLLIDFVXMMVKYGTISPEGYSYLENDEMIALOTNKLEKNATONISK-----L 354
Dd	421	EMKTQLESEDELQ----ASEDAKLRLVNNQALKGPQRDLQARDQEENEKRROLQROL 476
Qy	355	FPAPSEKSHBETDTKBEEAAKEYGSLKDSKTDNSNPCKTDEPK--GKTEAYLEAI 412
Dd	477	HEYETELEDENERALAAAACKLE-GDLKDLQLQADSIAKGREBAIKOLRKLQOMKDF 535
Qy	413	RKNJE-----WLKHDXKGNKEDYDLSKMRDFI-----NKOADAYVEKGILDK 455
Dd	536	QRELEDARASRDIEFATAKENEKKAKSILEADLMQLQEDLAAAAARAKQAD-----LEK 588
Qy	456	EE 457
Dd	589	EE 590
 RESULT 12 US-08-742-923A-6 ; Sequence 6, Application US/08742923A ; Patent No. 5869611 ; GENERAL INFORMATION: ; APPLICANT: Liu, Pu ; APPLICANT: Collins, Francis S. ; APPLICANT: Siciliano, Michael J. ; APPLICANT: Claxton, David ; TITLE OF INVENTION: Markers for Detection of Chromosome 16 ; TITLE OF INVENTION: Rearrangements ; NUMBER OF SEQUENCES: 14 ; CORRESPONDENCE ADDRESS: ; ADDRESSEE: Harness, Dickey & Pierce, P.L.C. ; STREET: P.O. Box 828 ; CITY: Bloomfield Hills ; STATE: MI ; COUNTRY: USA ; ZIP: 48303 ; COMPUTER READABLE FORM: ; MEDIUM TYPE: Floppy disk ; COMPUTER: IBM PC compatible ; OPERATING SYSTEM: PC-DOS/MS-DOS ; SOFTWARE: PatentIn Release #1.0, Version #1.25 ; CURRENT APPLICATION DATA: ; APPLICATION NUMBER: US/08/742,923A ; FILING DATE: No. 5869611member 1, 1996 ; CLASSIFICATION: 435 ; ATTORNEY/AGENT INFORMATION: ; NAME: Smith, Deann F. ; REGISTRATION NUMBER: 36683 ; REFERENCE/DOCKET NUMBER: 2115-00869DVC ; TELECOMMUNICATION INFORMATION: ; TELEPHONE: (810) 641-1600 ; TELEFAX: (810) 641-0270 ; INFORMATION FOR SEQ ID NO: 6: ; SEQUENCE CHARACTERISTICS: ; LENGTH: 816 amino acids ; TYPE: amino acid ; TOPOLOGY: linear		

MOLECULE TYPE: protein
US-08-742-923A-6

Query Match 6.6%; Score 159.5; DB 2; Length 816;
Best Local Similarity 22.6%; Pred. No. 0.0011; Indels 99; Gaps 21;

Matches 109; Conservative 81; Mismatches 193; Mismatches 190; Indels 99; Gaps 21;

```

Qy 26 SDDKSLHRELNSAERPLNEQIAEAEEDKIKKTPPENKPGQSNYSFV--DNLNLLKAITKE 83
Db 158 SHREEMENEVESVTCMLNE--AEGKAIKAKDVASLSQLODTQELLQEEETROKLNVTK 215
Qy 84 KEKTEKQSTR---SSPLDNKLNVEDVDSTKN-----RKLIDDYDSTKGLDH---KF 131
Db 216 LRQLEERNLSQDQDDEEMAKQNLERHISTLNQLSDSKKQLQDFASTVEALEEGKRFQKE 275
Qy 132 QDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFD-----KIVSKLLNLGLITESQA 186
Db 276 QKEIENLTQQYEEKAAAYDKLEKTKNRLOQELDLVVDLNDQRLVSNL-----EKKQ 328
Qy 187 HTLEDEVAELVQLKLSKEANNYEEDPNKPTSWTENQAGKIPKVTPTMA-AIODGLAKGEN 245
Db 329 RKPQLLAE--EKNIS---SKYADERDR---AEAEAREKETKALSILARALEEAKKEE 379
Qy 246 DETVSNLTTLNGLERTTYTSED-----NPEELQYFPNPFYALLKSIDSEKAKEKETLIT 301
Db 380 LERTNKMVK---AEMEDLVSSKDDVGVKNVHEL-----EKSRALETQME 420
Qy 302 IMKTLIDFVKMVKYGTISPGEVSVLENDEMIALQTKNLEKNATDNISK-----L 354
Db 421 EMKTLQEELELQ-----ASEDAKLRLEVNMQALKQGFERDLQARDEQNEEKRLQRLQ 476
Qy 355 FPAPSEKSHETDSTKEAAKMEKEYGSLKSDTKDSDNSNPGKTDDEPK--GKTEAYLEAI 412
Db 477 HEYETELEDERNERALLAAAKKLE-GDLKLELQADSIAIKGREAIAKQLKLAQMKDF 535
Qy 413 RKNIE-----WLKXHDKGNKEDYDLSKMRDPI-----NKQADAYVEKGILDK 455
Db 536 QRELEDAARDEIFAATAKENEKAKSLEADLMQLQEDLAAAEARAKQAD-----LEK 588
Qy 456 EE 457
Db 589 EE 590

```

RESULT 13

US-08-533-306A-4
Sequence 4, Application US/08533306A
Patent No. 5837457

GENERAL INFORMATION:
APPLICANT: Liu, Pu

APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.

APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16

TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828
CITY: Bloomfield Hills

STATE: MI
COUNTRY: USA

ZIP: 48303
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/533.306A

FILING DATE: September 25, 1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
COMPUTER: IBM PC compatible

NAME: Smith, DeAnn F.
REGISTRATION NUMBER: 36683
REFERENCE/DOCKET NUMBER: 2115-00869COB

TELEPHONE: (810) 641-1600
TELEFAX: (810) 641-0270

INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:

LENGTH: 885 amino acids
TYPE: amino acid

TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-533-306A-4

Query Match 6.6%; Score 158.5; DB 2; Length 885;
Best Local Similarity 22.9%; Pred. No. 0.0015;

Matches 109; Conservative 79; Mismatches 190; Indels 99; Gaps 21;

```

Qy 31 LHNRELSAERPLNEQIAEAEEDKIKKTPPENKPGQSNYSFV--DNLNLLKAITKEKIE 88
Db 232 LONEVESVTCMLNE--AEGKAIKAKDVASLSQLODTQELLQEEETROKLNVTKRLQLE 289
Qy 89 KEROSIR---SSPLDNKLNVEDVDSTKN-----RKLIDDYDSTKGLDH---KFQDDPD 136
Db 290 EERNLSQDQDDEEMAKQNLERHISTLNQLSDSKKQLQDFASTVEALEEGKRFQKEIE 349
Qy 137 GLHOLDGTPLTAEDIVHKIAARIYEENDRAVFD-----KIVSKLLNLGLITESQAHTLED 191
Db 350 NLTQQYEEKAAAYDKLEKTKNRLOQELDLVVDLNDQRLVSNL-----EKKQKFPQ 402
Qy 192 EVAELVQLKLSKEANNYEEDPNKPTSWTENQAGKIPKVTPTMA-AIODGLAKGENDETVS 250
Db 403 LLAE--EKNIS---SKYADERDR---AEAEAREKETKALSILARALEEAKKELETRN 453
Qy 251 NTLTLNGLERTTYTSED-----NPEELQYFPNPFYALLKSIDSEKAKEKETLITIMTKL 306
Db 454 KMLK---AEMEDLVSSKDDVGVKNVHEL-----EKSRALETQMEEMKTQ 494
Qy 307 IDFKMVKYGTISPGEVSVLENDEMIALQTKNLEKNATDNISK-----LFPAPS 359
Db 495 LEELEDELQ-----ASEDAKLRLEVNMQALKQGFERDLQARDEQNEEKRLQRLQHEYT 550
Qy 360 EKSHEETDSTKEAAKMEKEYGSLKSDTKDSDNSNPGKTDDEPK--GKTEAYLEAIRKNIE 417
Db 551 ELEDERNERALLAAAKKLE-GDLKLELQADSIAIKGREAIAKQLKLAQMKDFQRELE 609
Qy 418 -----WLKXHDKGNKEDYDLSKMRDPI-----NKQADAYVEKGILDKEE 457
Db 610 DARASDEIFAATAKENEKAKSLEADLMQLQEDLAAAEARAKQAD-----LEKEE 659

```

RESULT 14

US-08-742-923A-4

Sequence 4, Application US/08742923A
Patent No. 5869611

GENERAL INFORMATION:
APPLICANT: Liu, Pu

APPLICANT: Collins, Francis S.
APPLICANT: Siciliano, Michael J.

APPLICANT: Claxton, David
TITLE OF INVENTION: Markers for Detection of Chromosome 16

TITLE OF INVENTION: Rearrangements
NUMBER OF SEQUENCES: 14

CORRESPONDENCE ADDRESS:
ADDRESSEE: Harness, Dickey & Pierce, P.L.C.

STREET: P.O. Box 828
CITY: Bloomfield Hills

STATE: MI
COUNTRY: USA

ZIP: 48303
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/742.923A
; FILING DATE: No. 5869611member 1, 1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Smith, DeAnn F.
; REGISTRATION NUMBER: 36683
; REFERENCE/DOCKET NUMBER: 2115-00869DVC
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (810) 641-1600
; TELEFAX: (810) 641-0270
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 885 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-742-923A-4

Query Match      6.6%; Score 158.5; DB 2; Length 885;
Best Local Similarity 22.9%; Pred. No. 0.0015;
Matches 109; Conservative 79; Mismatches 190; Indels 99; Gaps 21;

Qy 31 LHNRELSAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSFV--DNLNLLKAITKEKIE 88
Db 232 LQNEVESVTGMLNE--AEQKAIKLAKDVASLSQLODTQELLQEBETROKLNVS TKLRQLE 289

Qy 89 KEROSIR---SSPLDNKLNVEDVSTKN-----RKLDDYDSTKSGLDH---KFQDDPD 136
Db 290 EERNLQDQDEMEAKONLHERHISTLNLIQSDSKKLQDFASTVALEEGKKRFQKEIE 349

Qy 137 GLHQLDGTPLTAEDIVHKIAAIEYENDRAVFD-----KIVSKLNLNGLITESQAHTLED 191
Db 350 NLTQYEEKAAAYDKLETKNRLOQLQELDDLVDLDNQRLVSNL-----EKKQKEDQ 402

Qy 192 EVAEVLQKLIKEANNVEEDPNKPTSWTENQAGKIPEKVTMA--AIQDGLAKGENDETVS 250
Db 403 LLAE--EKNIS---SKYADERDR---AAEAREKETKALSARALBEALEAKEELERTN 453

Qy 251 NTLTLTNGLERRTYSED---NFEELQYFPNFYALLKSIDSEKAKEKTLITIMKTL 306
Db 454 KQLK-----AEMEDLVSKDDVGKQVHQL-----BKSRALETQMEEMKTQ 494

Qy 307 IDFKVMVKYGTISPFGVSYLENLDEMIALQTKNLEKNATDNISK-----LFPAPS 359
Db 495 LEELEDELQ---ASEDAKLELVNMQALKQCFERDQARDEQNEEKRRQRLQRLHEYET 550

Qy 360 EKSHEETDSTKEAAKMEKEYSGLSKDSTKDDNSPGGKTDDEPK--GKTEAYLEAIRKNIE 417
Db 551 ELEDERNERALAAAKKLE--GDLKDLQADSIAIKGEEAIAKQLRKLQAKMDFQRELE 609

Qy 418 -----WLKHKDKGKNKEDYDLSKMRDFI-----NKQADAYVEKGILDKEE 457
Db 610 DARASRDEIFATAKENEKAKSLEADLMQLEDLAAAARAKQAD-----LEKEE 659
```

RESULT 15

```
US-09-914-259-12
; Sequence 12, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 12
```

```
; LENGTH: 1087
; TYPE: PRT
; ORGANISM: Oryctolagus cuniculus
; US-09-914-259-12

Query Match      6.5%; Score 156.5; DB 4; Length 1087;
Best Local Similarity 19.3%; Pred. No. 0.0028;
Matches 103; Conservative 83; Mismatches 180; Indels 167; Gaps 22;

Qy 31 LHNRELSAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSFVDNLNLLKAITKEKIEKE 90
Db 304 IHEKEVEIDR-LNEQIIKLOOQ-----LKITTDKNVIEEK 337

Qy 91 QRSIRSPPLDNKLNVEDVST-KNR-----KLIDDYDSTKSLDGHKFDDPDGLHQL 141
Db 338 NELIRDLEAQIECLMSQDERVRKNREBIEQLNEVIEKIQELANIQDKTSDPSSLSLSE- 396

Qy 142 DGTPLTAEDIVHKIAAIEYENDRAVFDKIVSKLNLNGLITESQAHTLEDVA----- 194
Db 397 -----EADSLKHQL-----DKVIAEKAL-----EHQVETTNNEEMAVTKNVLK 434

Qy 195 -----EVLQKLIS--KEANNVEEDPNKPTSWTENQAGKIPEKVTMAAI--ODGLAKG 243
Db 435 ETNPKMQLTQELCSLKREREKMERIQSVPEKSVNMSVGLSKDKPKMDLIPTDALLAQL 494

Qy 244 E-----NDETVSNTL-TLTNGLERRTKYSEDNFEELQYFPNFYALLKSIDSEKAKE 295
Db 495 ETOTQLRSSESSKVSLSLETLLQLESTVSTKDLLELTQCY-----KQIQDMREQGR 547

Qy 296 KET-----LITIMKTL-----IDFKVMVKYGTISPFG----- 324
Db 548 SETEMLTQKIVSLQKVLEEKVAAALVSQVQLAEVQYVVKLCADKPAVSSDPARTEVPGLS 607

Qy 325 -----VSYLENL-----DEMIALQTKNLEK-NATDNISKLF 355
Db 608 QLAGNTMESDVSAITWRISELESOLVEMHSLISEKEQVEIAEKNALEKEKQLQELQKLV 667

Qy 356 PAPSEKSHETDSTKEAAKMEKEYSGLSKDSTKDDNSNPG-----GKTDEPKGKTEAYL 409
Db 668 QDSETKQRE-----RERQSRLLHGLDGLVLESTTSEESGVFGELEALRAESAAPKGE 722

Qy 410 EATRKNIEWLKKHDKKGNKEDYDLSKMRDFINKQADA---YVEKGILDKEEAE 459
Db 723 ELAEKLOEELLVKETNNMASLPKLSHVDRDQTEAEDKLSHFSEKE--DKTEVQ 773

Search completed: January 2, 2004, 11:03:01
Job time : 23 secs
```

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 11:00:07 ; Search time 38 Seconds
(without alignments)
2471.055 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGWLVLPQAF.....EKGILDKKEAEAKRIYSSL 468

Scoring table:
Gapop 10.0 , Gapext 0.5

Searched: 733937 seqs, 200641211 residues

Total number of hits satisfying chosen parameters: 733937

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2399	100.0	468	11	US-09-931-836-67
2	2399	100.0	468	12	US-10-035-977-67
3	2399	100.0	468	12	US-10-063-735-150
4	2399	100.0	468	12	US-10-137-870-90
5	2399	100.0	468	12	US-10-140-018-90
6	2399	100.0	468	12	US-10-140-021-90
7	2399	100.0	468	12	US-10-140-274-90
8	2399	100.0	468	12	US-10-140-471-90
9	2399	100.0	468	12	US-10-140-807-90
10	2399	100.0	468	12	US-10-140-922-90
11	2399	100.0	468	12	US-10-140-924-90
12	2399	100.0	468	12	US-10-140-926-90
13	2399	100.0	468	12	US-10-141-698-90
14	2399	100.0	468	12	US-10-141-702-90
15	2399	100.0	468	12	US-10-141-704-90

16	2399	100.0	468	12	US-10-142-421-90	Sequence 90, Appl
17	2399	100.0	468	12	US-10-142-432-90	Sequence 90, Appl
18	2399	100.0	468	12	US-10-142-767-90	Sequence 90, Appl
19	2399	100.0	468	12	US-10-143-033-90	Sequence 90, Appl
20	2399	100.0	468	12	US-10-144-994-90	Sequence 90, Appl
21	2399	100.0	468	12	US-10-145-628-90	Sequence 90, Appl
22	2399	100.0	468	12	US-10-145-631-90	Sequence 90, Appl
23	2399	100.0	468	12	US-10-145-633-90	Sequence 90, Appl
24	2399	100.0	468	12	US-10-145-746-90	Sequence 90, Appl
25	2399	100.0	468	12	US-10-145-748-90	Sequence 90, Appl
26	2399	100.0	468	12	US-10-145-823-90	Sequence 90, Appl
27	2399	100.0	468	12	US-10-145-826-90	Sequence 90, Appl
28	2399	100.0	468	12	US-10-145-870-90	Sequence 90, Appl
29	2399	100.0	468	12	US-10-145-876-90	Sequence 90, Appl
30	2399	100.0	468	12	US-10-145-959-90	Sequence 90, Appl
31	2399	100.0	468	12	US-10-146-724-90	Sequence 90, Appl
32	2399	100.0	468	12	US-10-146-725-90	Sequence 90, Appl
33	2399	100.0	468	12	US-10-146-795-90	Sequence 90, Appl
34	2399	100.0	468	12	US-10-147-495-90	Sequence 90, Appl
35	2399	100.0	468	12	US-10-147-501-90	Sequence 90, Appl
36	2399	100.0	468	12	US-10-147-504-90	Sequence 90, Appl
37	2399	100.0	468	12	US-10-147-506-90	Sequence 90, Appl
38	2399	100.0	468	12	US-10-147-509-90	Sequence 90, Appl
39	2399	100.0	468	12	US-10-147-510-90	Sequence 90, Appl
40	2399	100.0	468	12	US-10-147-511-90	Sequence 90, Appl
41	2399	100.0	468	12	US-10-147-529-90	Sequence 90, Appl
42	2399	100.0	468	12	US-10-152-397-90	Sequence 90, Appl
43	2399	100.0	468	12	US-10-153-586-90	Sequence 90, Appl
44	2399	100.0	468	12	US-10-158-783-90	Sequence 90, Appl
45	2399	100.0	468	12	US-10-158-786-90	Sequence 90, Appl

ALIGNMENTS

RESULT 1

US-09-931-836-67
; Sequence 67, Application US/09931836
; Publication No. US20030027249A1
; GENERAL INFORMATION:
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Pan, James
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Watanabe, Colin K.
; APPLICANT: Wood, William I.
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3030R1C1
; CURRENT APPLICATION NUMBER: US/09/931,836
; CURRENT FILING DATE: 2001-08-16
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/112514
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113300
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113430
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/114140
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115552
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/116843
; PRIOR FILING DATE: 1999-01-22

; PRIOR APPLICATION NUMBER: 60/125774
 ; PRIOR FILING DATE: 1999-03-23
 ; PRIOR APPLICATION NUMBER: 60/125778
 ; PRIOR FILING DATE: 1999-03-23
 ; PRIOR APPLICATION NUMBER: 60/125826
 ; PRIOR FILING DATE: 1999-03-24
 ; PRIOR APPLICATION NUMBER: 60/127035
 ; PRIOR FILING DATE: 1999-03-31
 ; PRIOR APPLICATION NUMBER: 60/127706
 ; PRIOR FILING DATE: 1999-04-05
 ; PRIOR APPLICATION NUMBER: 60/129122
 ; PRIOR FILING DATE: 1999-04-13
 ; PRIOR APPLICATION NUMBER: 60/130359
 ; PRIOR FILING DATE: 1999-04-21
 ; PRIOR APPLICATION NUMBER: 60/131270
 ; PRIOR FILING DATE: 1999-04-27
 ; PRIOR APPLICATION NUMBER: 60/131272
 ; PRIOR FILING DATE: 1999-04-27
 ; PRIOR APPLICATION NUMBER: 60/131291
 ; PRIOR FILING DATE: 1999-04-27
 ; PRIOR APPLICATION NUMBER: 60/132371
 ; PRIOR FILING DATE: 1999-05-04
 ; PRIOR APPLICATION NUMBER: 60/132379
 ; PRIOR FILING DATE: 1999-05-04
 ; PRIOR APPLICATION NUMBER: 60/132383
 ; PRIOR FILING DATE: 1999-05-04
 ; PRIOR APPLICATION NUMBER: 60/135750
 ; PRIOR FILING DATE: 1999-05-25
 ; PRIOR APPLICATION NUMBER: 60/138166
 ; PRIOR FILING DATE: 1999-06-08
 ; PRIOR APPLICATION NUMBER: 60/144791
 ; PRIOR FILING DATE: 1999-07-20
 ; PRIOR APPLICATION NUMBER: 60/146970
 ; PRIOR FILING DATE: 1999-08-03
 ; PRIOR APPLICATION NUMBER: 60/162506
 ; PRIOR FILING DATE: 1999-10-29
 ; PRIOR APPLICATION NUMBER: 09/31182
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: 09/380142
 ; PRIOR FILING DATE: 1999-08-25
 ; PRIOR APPLICATION NUMBER: 09/644848
 ; PRIOR FILING DATE: 2000-08-22
 ; PRIOR APPLICATION NUMBER: 09/747259
 ; PRIOR FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: 09/816744
 ; PRIOR FILING DATE: 2001-03-22
 ; PRIOR APPLICATION NUMBER: 09/854208
 ; PRIOR FILING DATE: 2001-05-10
 ; PRIOR APPLICATION NUMBER: 09/854280
 ; PRIOR FILING DATE: 2001-05-10
 ; PRIOR APPLICATION NUMBER: 09/874503
 ; PRIOR FILING DATE: 2001-06-05
 ; PRIOR APPLICATION NUMBER: 09/869599
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: 09/908,827
 ; PRIOR FILING DATE: 2001-07-18
 ; PRIOR APPLICATION NUMBER: PCT/US99/10733
 ; PRIOR FILING DATE: 1999-05-14
 ; PRIOR APPLICATION NUMBER: PCT/US99/28551
 ; PRIOR FILING DATE: 1999-12-02
 ; PRIOR APPLICATION NUMBER: PCT/US99/30720
 ; PRIOR FILING DATE: 1999-12-22
 ; PRIOR APPLICATION NUMBER: PCT/US00/05601
 ; PRIOR FILING DATE: 2000-03-01
 ; PRIOR APPLICATION NUMBER: PCT/US00/05841
 ; PRIOR FILING DATE: 2000-03-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/14042
 ; PRIOR FILING DATE: 2000-05-22
 ; PRIOR APPLICATION NUMBER: PCT/US00/15264
 ; PRIOR FILING DATE: 2000-06-02
 ; PRIOR APPLICATION NUMBER: PCT/US00/23522
 ; PRIOR FILING DATE: 2000-08-23
 ; PRIOR APPLICATION NUMBER: PCT/US00/23328

; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: PCT/US00/32678
 ; PRIOR FILING DATE: 2000-12-01
 ; PRIOR APPLICATION NUMBER: PCT/US00/34956
 ; PRIOR FILING DATE: 2000-12-20
 ; PRIOR APPLICATION NUMBER: PCT/US01/06520
 ; PRIOR FILING DATE: 2001-08-28
 ; PRIOR APPLICATION NUMBER: PCT/US01/17800
 ; PRIOR FILING DATE: 2001-06-01
 ; PRIOR APPLICATION NUMBER: PCT/US01/19692
 ; PRIOR FILING DATE: 2001-06-20
 ; PRIOR APPLICATION NUMBER: PCT/US01/21066
 ; PRIOR FILING DATE: 2001-06-29
 ; PRIOR APPLICATION NUMBER: PCT/US01/21735
 ; PRIOR FILING DATE: 2001-07-09
 ; NUMBER OF SEQ ID NOS: 80
 ; SEQ ID NO 67
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 ; US-09-931-836-67
 Query Match 100.0%; Score 2399; DB 11; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8.7e-150;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MGFLGTGWTILVLPIQAFKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
 DB 1 MGFLGTGWTILVLPIQAFKPGGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
 QY 61 ENKPGQSNYSFVDNLNLLKAITEKIEKERSIRSSPLDNKLNVEDVDSTKNRKLIDDY 120
 DB 61 ENKPGQSNYSFVDNLNLLKAITEKIEKERSIRSSPLDNKLNVEDVDSTKNRKLIDDY 120
 QY 121 DSTKSGLDHKKFQDDPDGLHQLDGTPLTAEDI VHKAARIYEENDRAVFDKIVSKLLNLGL 180
 DB 121 DSTKSGLDHKKFQDDPDGLHQLDGTPLTAEDI VHKAARIYEENDRAVFDKIVSKLLNLGL 180
 QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKTSWTENQAGKIPEKVTMAAIQDGL 240
 DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKTSWTENQAGKIPEKVTMAAIQDGL 240
 QY 241 AKGENDETVSNTLTITNGLERRTKTSYSDNFEELQYFPNFVALLKSIDSEKEAKEKTLI 300
 DB 241 AKGENDETVSNTLTITNGLERRTKTSYSDNFEELQYFPNFVALLKSIDSEKEAKEKTLI 300
 QY 301 TIMKTLIDFVQMMVKYGTISP EEGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
 DB 301 TIMKTLIDFVQMMVKYGTISP EEGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
 QY 361 KSHEETDSTKEAAKMEKEYGSLKDDSTKDDNSPGKTDPEPKGTEAYLEAIRKNIEWLK 420
 DB 361 KSHEETDSTKEAAKMEKEYGSLKDDSTKDDNSPGKTDPEPKGTEAYLEAIRKNIEWLK 420
 QY 421 KHKKKGKEDYDLSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
 DB 421 KHKKKGKEDYDLSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
 RESULT 2
 US-10-035-977-67
 ; Sequence 67, Application US/10035977
 ; Publication No. US20030134327A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Eaton, Dan L.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Pan, James
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Watanabe, Colin K.
 ; APPLICANT: Wood, William I.

APPLICANT: Zhang Zemin
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3030R1C10
CURRENT APPLICATION NUMBER: US/10/035,977
CURRENT FILING DATE: 2001-12-26
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/112514
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113300
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113430
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/114140
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115552
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/116843
PRIOR FILING DATE: 1999-01-22
PRIOR APPLICATION NUMBER: 60/125774
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125778
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/125826
PRIOR FILING DATE: 1999-03-24
PRIOR APPLICATION NUMBER: 60/127035
PRIOR FILING DATE: 1999-03-31
PRIOR APPLICATION NUMBER: 60/127706
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/129122
PRIOR FILING DATE: 1999-04-13
PRIOR APPLICATION NUMBER: 60/130359
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131272
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/132371
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132379
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/132383
PRIOR FILING DATE: 1999-05-04
PRIOR APPLICATION NUMBER: 60/135750
PRIOR FILING DATE: 1999-05-25
PRIOR APPLICATION NUMBER: 60/138166
PRIOR FILING DATE: 1999-06-08
PRIOR APPLICATION NUMBER: 60/144791
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/146970
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/162506
PRIOR FILING DATE: 1999-10-29
PRIOR APPLICATION NUMBER: 09/311832
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 09/380142
PRIOR FILING DATE: 1999-08-25
PRIOR APPLICATION NUMBER: 09/644848
PRIOR FILING DATE: 2000-08-22
PRIOR APPLICATION NUMBER: 09/747259
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: 09/816744
PRIOR FILING DATE: 2001-03-22
PRIOR APPLICATION NUMBER: 09/854208
PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/854280

PRIOR FILING DATE: 2001-05-10
PRIOR APPLICATION NUMBER: 09/874503
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: 09/869599
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: 09/908,827
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: PCT/US99/10733
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: PCT/US99/28551
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30720
PRIOR FILING DATE: 1999-12-22
PRIOR APPLICATION NUMBER: PCT/US00/05601
PRIOR FILING DATE: 2000-03-01
PRIOR APPLICATION NUMBER: PCT/US00/05841
PRIOR FILING DATE: 2000-03-02
PRIOR APPLICATION NUMBER: PCT/US00/14042
PRIOR FILING DATE: 2000-05-22
PRIOR APPLICATION NUMBER: PCT/US00/15264
PRIOR FILING DATE: 2000-06-02
PRIOR APPLICATION NUMBER: PCT/US00/23522
PRIOR FILING DATE: 2000-08-23
PRIOR APPLICATION NUMBER: PCT/US00/23328
PRIOR FILING DATE: 2000-08-24
PRIOR APPLICATION NUMBER: PCT/US00/32678
PRIOR FILING DATE: 2000-12-01
PRIOR APPLICATION NUMBER: PCT/US00/34956
PRIOR FILING DATE: 2000-12-20
PRIOR APPLICATION NUMBER: PCT/US01/06520
PRIOR FILING DATE: 2001-02-28
PRIOR APPLICATION NUMBER: PCT/US01/17800
PRIOR FILING DATE: 2001-06-01
PRIOR APPLICATION NUMBER: PCT/US01/19692
PRIOR FILING DATE: 2001-06-20
PRIOR APPLICATION NUMBER: PCT/US01/21066
PRIOR FILING DATE: 2001-06-29
PRIOR APPLICATION NUMBER: PCT/US01/21735
PRIOR FILING DATE: 2001-07-09
NUMBER OF SEQ ID NOS: 80
SEQ ID NO 67
LENGTH: 468
TYPE: PRT
ORGANISM: Homo Sapien
US-10-035-977-67

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTVILVLVLPQAPPKPGSQDKSLHNRLSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTVILVLVLPQAPPKPGSQDKSLHNRLSAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVDNLNLLKAI TEKEKIERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
Db 61 ENKPGQSNYSFVDNLNLLKAI TEKEKIERQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
Qy 121 DSTKSGLDHDKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHDKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTTPMAAIODGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEDPNKPTSWTENQAGKPEKVTTPMAAIODGL 240
Qy 241 AKGENDETYSNTLTLTNGLERRTKTYSEDNFEBLQYFPNFPYALLKSIDSEKAKEKETLI 300
Db 241 AKGENDETYSNTLTLTNGLERRTKTYSEDNFEBLQYFPNFPYALLKSIDSEKAKEKETLI 300
Qy 301 TIMKTLIDFVKMMVKYGTGISPEEGSVYLENDLDEMIALQTKNLEKNATDNISKLFAPAPSE 360
Db 301 TIMKTLIDFVKMMVKYGTGISPEEGSVYLENDLDEMIALQTKNLEKNATDNISKLFAPAPSE 360

RESULT 5

US-10-140-018-90
 ; Sequence 90, Application US/10140018
 ; Publication No. US2003013885A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C158
 ; CURRENT APPLICATION NUMBER: US/10/140,018
 ; CURRENT FILING DATE: 2002-05-06
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 90
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-140-018-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8.7e-150;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGFLGTGTWILVLP	IOA	PPKPGSQD	KS	LHN	RELS	AE	RP	NE	QIA	AE	AE	DK	IK	TY	PP	60
Db	1	MGFLGTGTWILVLP	IOA	PPKPGSQD	KS	LHN	RELS	AE	RP	NE	QIA	AE	AE	DK	IK	TY	PP	60
Qy	61	ENKPGQSNYSFVDN	LNLL	KA	ITE	KE	KE	QSI	RS	SS	PL	NK	LV	ED	VD	ST	KN	KL
Db	61	ENKPGQSNYSFVDN	LNLL	KA	ITE	KE	KE	QSI	RS	SS	PL	NK	LV	ED	VD	ST	KN	KL
Qy	121	DSTKSGLDHFKFOD	PD	GL	HQ	LD	GT	PL	TA	ED	IV	HK	IA	RI	YE	ND	RA	VD
Db	121	DSTKSGLDHFKFOD	PD	GL	HQ	LD	GT	PL	TA	ED	IV	HK	IA	RI	YE	ND	RA	VD
Qy	181	ITESQAHTLEDEVA	EV	LQ	KL	IS	KE	AN	YE	ED	PN	KPT	SW	TE	NO	AG	KI	PE
Db	181	ITESQAHTLEDEVA	EV	LQ	KL	IS	KE	AN	YE	ED	PN	KPT	SW	TE	NO	AG	KI	PE
Qy	241	AKGENDET	VS	N	T	L	T	NG	L	R	R	T	K	T	Y	S	E	D
Db	241	AKGENDET	VS	N	T	L	T	NG	L	R	R	T	K	T	Y	S	E	D
Qy	301	TIMKT	L	I	D	F	V	K	M	V	K	Y	G	T	I	S	P	E
Db	301	TIMKT	L	I	D	F	V	K	M	V	K	Y	G	T	I	S	P	E
Qy	361	KSHEET	ST	KE	EA	K	ME	K	E	Y	G	S	L	K	D	ST	K	D
Db	361	KSHEET	ST	KE	EA	K	ME	K	E	Y	G	S	L	K	D	ST	K	D
Qy	421	KHDKKGN	K	E	D	Y	L	S	K	M	R	D	F	I	N	K	O	A
Db	421	KHDKKGN	K	E	D	Y	L	S	K	M	R	D	F	I	N	K	O	A

RESULT 6

US-10-140-274-90

US-10-140-021-90

; Sequence 90, Application US/10140021
 ; Publication No. US2003013886A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Baker, Kevin P.
 ; APPLICANT: Beresini, Maureen
 ; APPLICANT: Deforge, Laura
 ; APPLICANT: Desnoyers, Luc
 ; APPLICANT: Filvaroff, Ellen
 ; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C167
 ; CURRENT APPLICATION NUMBER: US/10/140,021
 ; CURRENT FILING DATE: 2002-05-06
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 90
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-140-021-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8.7e-150;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGFLGTGTWILVLP	IOA	PPKPGSQD	KS	LHN	RELS	AE	RP	NE	QIA	AE	AE	DK	IK	TY	PP	60
Db	1	MGFLGTGTWILVLP	IOA	PPKPGSQD	KS	LHN	RELS	AE	RP	NE	QIA	AE	AE	DK	IK	TY	PP	60
Qy	61	ENKPGQSNYSFVDN	LNLL	KA	ITE	KE	KE	QSI	RS	SS	PL	NK	LV	ED	VD	ST	KN	KL
Db	61	ENKPGQSNYSFVDN	LNLL	KA	ITE	KE	KE	QSI	RS	SS	PL	NK	LV	ED	VD	ST	KN	KL
Qy	121	DSTKSGLDHFKFOD	PD	GL	HQ	LD	GT	PL	TA	ED	IV	HK	IA	RI	YE	ND	RA	VD
Db	121	DSTKSGLDHFKFOD	PD	GL	HQ	LD	GT	PL	TA	ED	IV	HK	IA	RI	YE	ND	RA	VD
Qy	181	ITESQAHTLEDEVA	EV	LQ	KL	IS	KE	AN	YE	ED	PN	KPT	SW	TE	NO	AG	KI	PE
Db	181	ITESQAHTLEDEVA	EV	LQ	KL	IS	KE	AN	YE	ED	PN	KPT	SW	TE	NO	AG	KI	PE
Qy	241	AKGENDET	VS	N	T	L	T	NG	L	R	R	T	K	T	Y	S	E	D
Db	241	AKGENDET	VS	N	T	L	T	NG	L	R	R	T	K	T	Y	S	E	D
Qy	301	TIMKT	L	I	D	F	V	K	M	V	K	Y	G	T	I	S	P	E
Db	301	TIMKT	L	I	D	F	V	K	M	V	K	Y	G	T	I	S	P	E
Qy	361	KSHEET	ST	KE	EA	K	ME	K	E	Y	G	S	L	K	D	ST	K	D
Db	361	KSHEET	ST	KE	EA	K	ME	K	E	Y	G	S	L	K	D	ST	K	D
Qy	421	KHDKKGN	K	E	D	Y	L	S	K	M	R	D	F	I	N	K	O	A
Db	421	KHDKKGN	K	E	D	Y	L	S	K	M	R	D	F	I	N	K	O	A

; Sequence 90, Application US/10140274
; Publication No. US20030143674A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C161
; CURRENT APPLICATION NUMBER: US/10/140,274
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-274-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGTLVILVLPPIQAPFKPGSQDQKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
DB 1 MGFLGTGTLVILVLPPIQAPFKPGSQDQKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
QY 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
DB 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
QY 121 DSTKSGLDHFKFQDDPGHLQDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
DB 121 DSTKSGLDHFKFQDDPGHLQDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
QY 241 AKGENDTVSNTLTNLGLRRTKTYSEDNPEELQYFPNPFVALLKSIDSEKEAKEKETLI 300
DB 241 AKGENDTVSNTLTNLGLRRTKTYSEDNPEELQYFPNPFVALLKSIDSEKEAKEKETLI 300
QY 301 TIMKTLIDFVKMVKYCTISPEEGSVYLENIDEMIALQTKNLEKNATDNISKLFPAPSE 360
DB 301 TIMKTLIDFVKMVKYCTISPEEGSVYLENIDEMIALQTKNLEKNATDNISKLFPAPSE 360
QY 361 KSHEETDSTKEAAKMEKEYGSLKSDTSDNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
DB 361 KSHEETDSTKEAAKMEKEYGSLKSDTSDNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
QY 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
DB 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468

RESULT 8
US-10-140-471-90
; Sequence 90, Application US/10140471

; Publication No. US20030138887A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P330R1C163
; CURRENT APPLICATION NUMBER: US/10/140,471
; CURRENT FILING DATE: 2002-05-06
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-471-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGFLGTGTLVILVLPPIQAPFKPGSQDQKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
DB 1 MGFLGTGTLVILVLPPIQAPFKPGSQDQKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
QY 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
DB 61 ENKPGQSNYSFVDNLLKKAITEKEKIEKERSQSRSSPLDNKLNVEDVDSTKNRKLIDDY 120
QY 121 DSTKSGLDHFKFQDDPGHLQDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
DB 121 DSTKSGLDHFKFQDDPGHLQDGTPLTAEDIVHKIAARIYEENDRAVFDKIVSKLLNLGL 180
QY 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
DB 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIODGL 240
QY 241 AKGENDTVSNTLTNLGLRRTKTYSEDNPEELQYFPNPFVALLKSIDSEKEAKEKETLI 300
DB 241 AKGENDTVSNTLTNLGLRRTKTYSEDNPEELQYFPNPFVALLKSIDSEKEAKEKETLI 300
QY 301 TIMKTLIDFVKMVKYCTISPEEGSVYLENIDEMIALQTKNLEKNATDNISKLFPAPSE 360
DB 301 TIMKTLIDFVKMVKYCTISPEEGSVYLENIDEMIALQTKNLEKNATDNISKLFPAPSE 360
QY 361 KSHEETDSTKEAAKMEKEYGSLKSDTSDNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
DB 361 KSHEETDSTKEAAKMEKEYGSLKSDTSDNSPGGKTDEPKGTEAYLEAIRKNIEWLK 420
QY 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468
DB 421 KHDKGNKEDYDLSKMRDFINKQADAYVEKGILDKEEAEAIKRIYSSL 468

RESULT 9
US-10-140-807-90
; Sequence 90, Application US/10140807
; Publication No. US20030134354A1

```

; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C174
; CURRENT APPLICATION NUMBER: US/10/140,807
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-807-90

```

```

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTLVLVLPQAPFKPGSQDKSLHNRSLAERPLNEQIAEAEEDKIKTYPP 60
Db 1 MGFLGTGTLVLVLPQAPFKPGSQDKSLHNRSLAERPLNEQIAEAEEDKIKTYPP 60
Qy 61 ENKPGQSNYSFVDNLNLLKAI TEKEIEKERQSI RSPDLNKNLVNEDVDSTKRNKLIDDY 120
Db 61 ENKPGQSNYSFVDNLNLLKAI TEKEIEKERQSI RSPDLNKNLVNEDVDSTKRNKLIDDY 120
Qy 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDI VHKAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDI VHKAARIYEENDRAVDFKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTTPMAAIQDGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTTPMAAIQDGL 240
Qy 241 AKGENDETVSNTLTITNGLERRTKTYSEDNFEELOYPFNFYALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETVSNTLTITNGLERRTKTYSEDNFEELOYPFNFYALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLDIFVKMVKYGTISP EEGSVYLENDEMIALQTKNKLKNATDNISKLPAPSE 360
Db 301 TIMKTLDIFVKMVKYGTISP EEGSVYLENDEMIALQTKNKLKNATDNISKLPAPSE 360
Qy 361 KSHETDSTKEAAKMEKEYGSLKSDTKDNNPCKGTDPEPKGTEAYLEAIRKNIEWLK 420
Db 361 KSHETDSTKEAAKMEKEYGSLKSDTKDNNPCKGTDPEPKGTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKNGEDYDLSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
Db 421 KHKKGKNGEDYDLSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468

```

```

RESULT 10
US-10-140-922-90
; Sequence 90, Application US/10140922
; Publication No. US2003013889A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

```

; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: Deforge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C179
; CURRENT APPLICATION NUMBER: US/10/140,922
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-922-90

```

```

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTLVLVLPQAPFKPGSQDKSLHNRSLAERPLNEQIAEAEEDKIKTYPP 60
Db 1 MGFLGTGTLVLVLPQAPFKPGSQDKSLHNRSLAERPLNEQIAEAEEDKIKTYPP 60
Qy 61 ENKPGQSNYSFVDNLNLLKAI TEKEIEKERQSI RSPDLNKNLVNEDVDSTKRNKLIDDY 120
Db 61 ENKPGQSNYSFVDNLNLLKAI TEKEIEKERQSI RSPDLNKNLVNEDVDSTKRNKLIDDY 120
Qy 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDI VHKAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDI VHKAARIYEENDRAVDFKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTTPMAAIQDGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTTPMAAIQDGL 240
Qy 241 AKGENDETVSNTLTITNGLERRTKTYSEDNFEELOYPFNFYALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETVSNTLTITNGLERRTKTYSEDNFEELOYPFNFYALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLDIFVKMVKYGTISP EEGSVYLENDEMIALQTKNKLKNATDNISKLPAPSE 360
Db 301 TIMKTLDIFVKMVKYGTISP EEGSVYLENDEMIALQTKNKLKNATDNISKLPAPSE 360
Qy 361 KSHETDSTKEAAKMEKEYGSLKSDTKDNNPCKGTDPEPKGTEAYLEAIRKNIEWLK 420
Db 361 KSHETDSTKEAAKMEKEYGSLKSDTKDNNPCKGTDPEPKGTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKNGEDYDLSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468
Db 421 KHKKGKNGEDYDLSKMRDFINKQADAYVEKGILDKKEAEAIKRIYSSL 468

```

```

RESULT 11
US-10-140-924-90
; Sequence 90, Application US/10140924
; Publication No. US20030134355A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.

```

```

; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C177
; CURRENT APPLICATION NUMBER: US/10/140,924
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-924-90

```

```

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTTWLVLPQAPKPGSQDSKLSHNSLSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTTWLVLPQAPKPGSQDSKLSHNSLSAERPLNEQIAEAEEDKIKKTYPP 60

Qy 61 ENKPGQSNYSFVDNLLKKAITEKEIEKERQSISSPLDNKLVNVEDVSTKRLIDY 120
Db 61 ENKPGQSNYSFVDNLLKKAITEKEIEKERQSISSPLDNKLVNVEDVSTKRLIDY 120

Qy 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180

Qy 181 ITESQATLDEVAEVLQKLISKEANNYEDPNKPTSWTENQAGKIPKVTMPMAIQDGL 240
Db 181 ITESQATLDEVAEVLQKLISKEANNYEDPNKPTSWTENQAGKIPKVTMPMAIQDGL 240

Qy 241 AKGNDSTVNTLTTLTNGLRRTKTSYSEDNFEELQYFPNFPYALLKSIDSEKAEKETLI 300
Db 241 AKGNDSTVNTLTTLTNGLRRTKTSYSEDNFEELQYFPNFPYALLKSIDSEKAEKETLI 300

Qy 301 TIMKTLIDFVMMVKYGTISPESGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVMMVKYGTISPESGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360

Qy 361 KSHEETDSTKEEAAKEKEYGSLKSDTKDONSPPGKTDPEKGTAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEEAAKEKEYGSLKSDTKDONSPPGKTDPEKGTAYLEAIRKNIEWLK 420

Qy 421 KHKDKGNKEDYDLSKMRDFINKQADAVVEKILDKKEAEAIKRIYSSL 468
Db 421 KHKDKGNKEDYDLSKMRDFINKQADAVVEKILDKKEAEAIKRIYSSL 468

```

```

RESULT 12
US-10-140-926-90
; Sequence 90, Application US/10140926
; Publication No. US20030134356A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen

```

```

; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C187
; CURRENT APPLICATION NUMBER: US/10/140,926
; CURRENT FILING DATE: 2002-05-07
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-140-926-90

```

```

Query Match 100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8,7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTTWLVLPQAPKPGSQDSKLSHNSLSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTTWLVLPQAPKPGSQDSKLSHNSLSAERPLNEQIAEAEEDKIKKTYPP 60

Qy 61 ENKPGQSNYSFVDNLLKKAITEKEIEKERQSISSPLDNKLVNVEDVSTKRLIDY 120
Db 61 ENKPGQSNYSFVDNLLKKAITEKEIEKERQSISSPLDNKLVNVEDVSTKRLIDY 120

Qy 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180

Qy 181 ITESQATLDEVAEVLQKLISKEANNYEDPNKPTSWTENQAGKIPKVTMPMAIQDGL 240
Db 181 ITESQATLDEVAEVLQKLISKEANNYEDPNKPTSWTENQAGKIPKVTMPMAIQDGL 240

Qy 241 AKGNDSTVNTLTTLTNGLRRTKTSYSEDNFEELQYFPNFPYALLKSIDSEKAEKETLI 300
Db 241 AKGNDSTVNTLTTLTNGLRRTKTSYSEDNFEELQYFPNFPYALLKSIDSEKAEKETLI 300

Qy 301 TIMKTLIDFVMMVKYGTISPESGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360
Db 301 TIMKTLIDFVMMVKYGTISPESGVSYLENDEMIALQTKNLEKNATDNISKLPAPSE 360

Qy 361 KSHEETDSTKEEAAKEKEYGSLKSDTKDONSPPGKTDPEKGTAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEEAAKEKEYGSLKSDTKDONSPPGKTDPEKGTAYLEAIRKNIEWLK 420

Qy 421 KHKDKGNKEDYDLSKMRDFINKQADAVVEKILDKKEAEAIKRIYSSL 468
Db 421 KHKDKGNKEDYDLSKMRDFINKQADAVVEKILDKKEAEAIKRIYSSL 468

```

```

RESULT 13
US-10-141-698-90
; Sequence 90, Application US/10141698
; Publication No. US20030134357A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura

```

```

; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C206
; CURRENT APPLICATION NUMBER: US/10/141,698
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-698-90

Query Match      100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVNDNLNLLKAITKEKIEKERSQSRSSPLDNKNVEDVDSKRNKLLIDY 120
Db 61 ENKPGQSNYSFVNDNLNLLKAITKEKIEKERSQSRSSPLDNKNVEDVDSKRNKLLIDY 120
Qy 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTPTMAAIQDGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTPTMAAIQDGL 240
Qy 241 AKGENDETVSNTLTITNGLERTTKTYSNDFEELQYFNFYALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETVSNTLTITNGLERTTKTYSNDFEELQYFNFYALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLIDFVKMVKYGTISPBEVSVYLENDEMIALQTKNLEKNATDNISKLPFAPSE 360
Db 301 TIMKTLIDFVKMVKYGTISPBEVSVYLENDEMIALQTKNLEKNATDNISKLPFAPSE 360
Qy 361 KSHEETDSTKEAAKMEKEYGSLKSDTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEAAKMEKEYGSLKSDTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKKNEDYDLSKMRDFINKQADAYVEKGILDKEEAIAIKRIYSSL 468
Db 421 KHKKGKKNEDYDLSKMRDFINKQADAYVEKGILDKEEAIAIKRIYSSL 468

RESULT 14
US-10-141-702-90
; Sequence 90, Application US/10141702
; Publication No. US20030134358A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc

```

```

; APPLICANT: Filvaroff, Ellen
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Sherwood, Steven
; APPLICANT: Smith, Victoria
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tamas, Daniel
; APPLICANT: Watanabe, Colin K
; APPLICANT: Wood, William
; APPLICANT: Zhang, Zemin
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3330R1C208
; CURRENT APPLICATION NUMBER: US/10/141,702
; CURRENT FILING DATE: 2002-05-08
; Prior Application removed - See Palm or File Wrapper
; NUMBER OF SEQ ID NOS: 550
; SEQ ID NO 90
; LENGTH: 468
; TYPE: PRT
; ORGANISM: Homo Sapien
US-10-141-702-90

Query Match      100.0%; Score 2399; DB 12; Length 468;
Best Local Similarity 100.0%; Pred. No. 8.7e-150;
Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGFLGTGTTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Db 1 MGFLGTGTTWILVLPIQAFPKPGSQDKSLHNRELSAERPLNEQIAEAEEDKIKKTYPP 60
Qy 61 ENKPGQSNYSFVNDNLNLLKAITKEKIEKERSQSRSSPLDNKNVEDVDSKRNKLLIDY 120
Db 61 ENKPGQSNYSFVNDNLNLLKAITKEKIEKERSQSRSSPLDNKNVEDVDSKRNKLLIDY 120
Qy 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Db 121 DSTKSGLDHFKQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL 180
Qy 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTPTMAAIQDGL 240
Db 181 ITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVTPTMAAIQDGL 240
Qy 241 AKGENDETVSNTLTITNGLERTTKTYSNDFEELQYFNFYALLKSIDSEKEAKEKETLI 300
Db 241 AKGENDETVSNTLTITNGLERTTKTYSNDFEELQYFNFYALLKSIDSEKEAKEKETLI 300
Qy 301 TIMKTLIDFVKMVKYGTISPBEVSVYLENDEMIALQTKNLEKNATDNISKLPFAPSE 360
Db 301 TIMKTLIDFVKMVKYGTISPBEVSVYLENDEMIALQTKNLEKNATDNISKLPFAPSE 360
Qy 361 KSHEETDSTKEAAKMEKEYGSLKSDTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLK 420
Db 361 KSHEETDSTKEAAKMEKEYGSLKSDTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLK 420
Qy 421 KHKKGKKNEDYDLSKMRDFINKQADAYVEKGILDKEEAIAIKRIYSSL 468
Db 421 KHKKGKKNEDYDLSKMRDFINKQADAYVEKGILDKEEAIAIKRIYSSL 468

RESULT 15
US-10-141-704-90
; Sequence 90, Application US/10141704
; Publication No. US20030134359A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Beresini, Maureen
; APPLICANT: DeForge, Laura
; APPLICANT: Desnoyers, Luc
; APPLICANT: Filvaroff, Ellen

```

; APPLICANT: Gao, Wei-Qiang
 ; APPLICANT: Gerritsen, Mary E.
 ; APPLICANT: Goddard, Audrey
 ; APPLICANT: Godowski, Paul J.
 ; APPLICANT: Gurney, Austin L.
 ; APPLICANT: Sherwood, Steven
 ; APPLICANT: Smith, Victoria
 ; APPLICANT: Stewart, Timothy A.
 ; APPLICANT: Tumas, Daniel
 ; APPLICANT: Watanabe, Colin K
 ; APPLICANT: Wood, William
 ; APPLICANT: Zhang, Zemin
 ; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
 ; FILE REFERENCE: P3330R1C209
 ; CURRENT APPLICATION NUMBER: US/10/141,704
 ; CURRENT FILING DATE: 2002-05-08
 ; Prior Application removed - See Palm or File Wrapper
 ; NUMBER OF SEQ ID NOS: 550
 ; SEQ ID NO 90
 ; LENGTH: 468
 ; TYPE: PRT
 ; ORGANISM: Homo Sapien
 US-10-141-704-90

Query Match 100.0%; Score 2399; DB 12; Length 468;
 Best Local Similarity 100.0%; Pred. No. 8,7e-150;
 Matches 468; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	MGFLGTGTLVLVLPTQAFPKPGSQDKSLHRELNSAERPLNEQIAEAEEDKIKKTYPP	60
Db	1	MGFLGTGTLVLVLPTQAFPKPGSQDKSLHRELNSAERPLNEQIAEAEEDKIKKTYPP	60
Qy	61	ENKPGQSNYSFVDNLNLLKATEKEKEKQSRSSPLDNKLNVEDVDSTKNRKLDDY	120
Db	61	ENKPGQSNYSFVDNLNLLKATEKEKEKQSRSSPLDNKLNVEDVDSTKNRKLDDY	120
Qy	121	DSTKSGLDHFKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL	180
Db	121	DSTKSGLDHFKFQDDPDGLHQLDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLLNLGL	180
Qy	181	ITESQAHTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIQDGL	240
Db	181	ITESQAHTLEDEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPKVTMPMAAIQDGL	240
Qy	241	AKGENDETIVNTLTNGLERRTKYSEDNFEELQYFPNFYALUKSIDSEKEAKEKTLI	300
Db	241	AKGENDETIVNTLTNGLERRTKYSEDNFEELQYFPNFYALUKSIDSEKEAKEKTLI	300
Qy	301	TIMKTLIDFVKMVKYGTISPESGVSYLENLDEMIALQTKNLEKNATDNISKLPAPSE	360
Db	301	TIMKTLIDFVKMVKYGTISPESGVSYLENLDEMIALQTKNLEKNATDNISKLPAPSE	360
Qy	361	KSHETDSTKEEAKMEKEYGSLKSDKDNNSPGGKTDPEPKGTAYLEAIRKNIEWLK	420
Db	361	KSHETDSTKEEAKMEKEYGSLKSDKDNNSPGGKTDPEPKGTAYLEAIRKNIEWLK	420
Qy	421	KHDKGNKEDYDLSKMRDFINKQADAVEKGILDKEEAEAIKRIYSSL	468
Db	421	KHDKGNKEDYDLSKMRDFINKQADAVEKGILDKEEAEAIKRIYSSL	468

Search completed: January 2, 2004, 11:03:51
 Job time : 39 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 10:59:27 ; Search time 21 Seconds
(without alignments)
2143.188 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGTLVLVLPQAF.....EKGLDKERAEAIKRIYSSL 468
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 283308

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 76.*
1: pir1.*
2: pir2.*
3: pir3.*
4: pir4.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1957	81.6	533	2 A37180	chromogranin/secre
2	196	8.2	2139	2 T18296	myosin heavy chain
3	191	8.0	2454	1 QRMSP1	microtubule-associ
4	186.5	7.8	2116	2 A26655	myosin heavy chain
5	184	7.7	665	2 B71609	hypothetical prote
6	183.5	7.6	1790	2 S67593	transport protein
7	180.5	7.5	2364	2 A56577	microtubule-associ
8	179.5	7.5	1526	2 A45605	mature-parasite-in
9	179	7.5	853	2 T51505	hypothetical prote
10	179	7.5	1005	2 A64465	hypothetical prote
11	179	7.5	1558	2 B71603	RESA-H3 antigen PF
12	179	7.5	2269	2 T28677	rhostry protein -
13	178.5	7.4	1392	2 A43336	microtubule-vesicl
14	178.5	7.4	1427	2 S22695	restin - human
15	177	7.4	1871	2 D96796	probable heat shoc
16	172.5	7.2	1134	2 A60234	Iga Fc receptor pr
17	172.5	7.2	1164	1 FCSOAG	Iga Fc receptor pr
18	172.5	7.2	1979	2 C71622	hypothetical prote
19	171.5	7.1	3488	2 T34418	hypothetical prote
20	170	7.1	1804	2 T34518	nestin golden ha
21	169	7.0	1272	2 C90593	hypothetical prote
22	168.5	7.0	746	2 T47237	myosin II heavy ch
23	168.5	7.0	926	2 T24923	hypothetical prote
24	168	7.0	2401	2 T28676	rhostry protein -
25	167	7.0	976	2 T01553	hypothetical prote
26	166.5	6.9	1875	2 S38173	myosin-like protei
27	166.5	6.9	5327	2 T13564	microtubule-associ
28	166	6.9	991	2 H86168	hypothetical prote
29	166	6.9	1192	2 A71623	probable secreted

30	166	6.9	2663	1 S28261	centromere protein
31	164.5	6.9	719	2 A81358	hypothetical prote
32	164.5	6.9	911	2 S51441	hypothetical prote
33	164.5	6.9	1127	2 T28317	ORF MSV156 hypothe
34	164.5	6.9	1269	2 F84730	probable myosin he
35	164	6.8	1909	2 A45592	liver stage antige
36	163	6.8	1252	2 B42771	reticulocyte-bindi
37	163	6.8	2510	2 T28160	hypothetical prote
38	162.5	6.8	1738	2 T14967	interaptin - slime
39	162.5	6.8	2385	2 A32491	myosin heavy chain
40	162.5	6.8	2411	2 B32491	myosin heavy chain
41	161.5	6.7	1302	1 JC6009	surface-located me
42	161.5	6.7	1992	1 S02771	myosin heavy chain
43	161	6.7	852	2 D72230	conserved hypothet
44	160	6.7	1354	2 S74244	serine/threonine-s
45	159.5	6.6	821	2 S67087	hypothetical prote

ALIGNMENTS

RESULT 1

A37180 chromogranin/secretogranin-like vesicle protein precursor - rat

C;Species: Rattus norvegicus (Norway rat)

C;Date: 07-Feb-1992 #sequence_revision 07-Feb-1992 #text_change 30-Sep-1993

C;Accession: A37180

R;Otterger, H.P.; Battenberg, E.F.; Tsou, A.P.; Bloom, F.E.; Sutcliffe, J.G.

J. Neurosci. 10, 3135-3147, 1990

A;Title: IB1075: a brain- and pituitary-specific mRNA that encodes a novel chromogranin/

A;Reference number: A37180; MUID:90376160; PMID:2204688

A;Accession: A37180

A;Status: preliminary; not compared with conceptual translation

A;Molecule type: mRNA

A;Residues: 1-533 <OTT>

Query Match 81.6%; Score 1957; DB 2; Length 533;

Best Local Similarity 86.7%; Pred. No. 3.5e-87;

Matches 383; Conservative 21; Mismatches 38; Indels 0; Gaps 0;

QY	16	PTQAFPKPGSQKSLHNRELSAERPLNEQIAEAEADKIKKTPPENKPGQSNVSVFVNL	75
DB	63	PNSFPKPGSQKSLHNRELSAERPLNEQIAEAEADKIKKTPPENKPGQSNVSVFVNL	122
QY	76	NLLKATTEKIEKIERQSRSSPLDNKLVNVEDVDSTKNRKLIDDYDSTKSLDHFODDP	135
DB	123	NLLKATTEKIEKIERQSRSSPLDNKLVNVEDVDSTKNRKLIDDYDSTKSLDHFODDP	182
QY	136	DGLHQLDGTPLTAEDIVHKAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE	195
DB	183	DGLHQLDGTPLTAEDIVHKAARIYEENDRAVFDKIVSKLLNLGLITESQAHTLEDEVAE	242
QY	196	VLOKLISKANNVEEDPNKPTSWENOAGKIPKVTPTMAIQDGLAKGENDETIVSNLTLL	255
DB	243	ALQKLISKANNVEEDPNKPTSWENOAGKIPKVTPTMAIQDGLAKGENDETIVSNLTLL	302
QY	256	TNGLERTKTYSEDNEFEELQYFNFYALLKSIDSEKEAKEKETLITIMTKLIDFVKMVK	315
DB	303	SNGLERTKTYSEDNEFEELQYFNFYALLKSIDSEKEAKEKETLITIMTKLIDFVKMVK	362
QY	316	YGTISPPEGVSVLENIDEMIALQTKNLEQNATDNISKLPAPSEKSHETDSTKEAAK	375
DB	363	YGTISPPEGVSVLENIDEMIALQTKNLEQNATDNISKLPAPSEKSHETDSTKEAAK	422
QY	376	MEKEYGSLKDKSTKDDNSNPGGKTDEPKGTAEYLAIRKNIEWLKKHKDKGNKEDYDLK	435
DB	423	MEKEYGSLKDKSTKDDNSNPGGKTDEPKGTAEYLAIRKNIEWLKKHKDKGNKEDYDLK	482
QY	436	MRDFPINKQADAYVEKGILQKEE	457
DB	483	MRDFPINKQADAYVEKGILIRKK	504

A:Residues: 1-2464 <NOB>
A:Cross-references: EMBL:X51396; NID:G52999; PIDN:CAA3761.1; PID:G53000
R:Sanchez, C.; Padilla, R.; Paciucci, R.; Zabala, J.C.; Avila, J.
Arch. Biochem. Biophys. 310, 428-432, 1994
A>Title: Binding of heat-shock protein 70 (hap70) to tubulin.
F:589-592, 639-642, 649-652, 655-658, 660-663, 668-671, 674-677, 679-682, 683-686, 687-690, 691-694
F:1953/Binding site: phosphate (Tyr)
A:Accession: S44387 MUID:94234720; PMID:8179328
A>Status: preliminary
A:Molecule type: protein
A:Residues: 853-663, IC<SAN>
C:Superfamily: microtubule-associated protein MAP1B
C:Keywords: microtubule binding; phosphoprotein; tandem repeat
F:589-786/Domain: microtubule binding #status experimental <MTB>
F:589-592, 639-642, 649-652, 655-658, 660-663, 668-671, 674-677, 679-682, 683-686, 687-690, 691-694
R:K-E/D-XI
F:1861-2064/Region: 17-residue repeats
F:91,116,351,888,1124,1153,1169,1308,1662,1877,1918,2003,2030,2054,2083/Binding site: phosphate (Thr) (cov)
F:147,969,1336,1562,1563,1702,1708,1990,2057,2063,2419/Binding site: phosphate (Thr) (cov)
F:1953/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 8.0%; Score 191; DB 1; Length 2464;
Best Local Similarity 21.3%; Pred. No. 0.21;
Matches 118; Conservative 85; Mismatches 186; Indels 166; Gaps 24;

QY 22 KPGSQDSKLHRELSEARPLNEQIAEA-----EEDKKIKTYP--PENKPGOSNYSFVD 73
||| : : : : : : : : : : : : : : : : :
DB 550 KPVAASKVRKESKEETPEVTKTSQVEKTPRKVESKEVLVKDKPKVKTESKP ----- 600

QY 74 NLMLLKAITKEKTEKERQSIRSPDLNKLNVDDSTKNRKLIDDYDSTKSLGHKFQD 133
||| : : : : : : : : : : : : : : : : :
DB 601 -----SVTEKVVSKBEQSPVKAIEAQATESPVKTKDVVKKEIKTK--LEEKKEE 652

QY 134 DP-DGLHQLDGTPL-----TAEDIVHKIARIYEENDRAVFDPKIVSKLLNLGLITESQ 185
||| : : : : : : : : : : : : : : : : :
DB 653 KPKEVVKVEDKTPLKXDEKPRKEEVKKEIKKEIKKE-ERKELKKEVKK-----ETP 703

QY 186 AHTLEDEVAEVLQKLISKEANNYSEDPNK-----PTSWTENQAGKIPE--- 228
||| : : : : : : : : : : : : : : : : :
DB 704 LKDAKKEVKKEEKEVKKE-----EKEPKKEIKKISKDIKKTPOSDTKPKPSALKPRVAKK 759

QY 229 ----KVTPMAA--IQD-GLAK-----GENDETIVSNLT----- 253
DB 760 EESTKKEPLAAGKLKDQGVKVIKKEGKTTEAAATAVGTAATTAAVVAAAAGIASGPVKE 819

QY 254 ----TLTNGLERRTKTYSENFBELQVPFPNYALLKSIDSEKAKE-----KET 298
||| : : : : : : : : : : : : : : : : :
DB 820 LEAERSLMSSPEDTKOPEELKAEIDVAKOIKPOLIELIDEKLEKTPQGEAYVIOKET 879

QY 299 LIITMKTLLDFVMWVYKGTISPESGVSYLENDEMIALQTGNKL---EKNDATDNISKL- 354
||| : : : : : : : : : : : : : : : : :
DB 880 EVS-----KGSAREPDGGITTTTEGEQE--CEQTPEELPVEKQGVDDIEKFE 924

QY 355 ----FPAPSEKSHBEETDSTKEEAARKVEKEYSLKDDSTKDDNSNPQGKTDPKPKTEAYL 409
||| : : : : : : : : : : : : : : : : :
DB 925 DEGAGFSESSETGDVEEKAETEAEPEEPD-GEDNAGSGASAKHSPTEDDESASAADVHL 983

QY 410 EAIRNQI-----EWLKK-----HDKKGNKEDY-----DLSKMWRFIN 441
||| : : : : : : : : : : : : : : : : :
DB 984 KEKREVSVDGRABEDMDVLEKGAEQSEBEGBEDKAEADAREEGYEPDKTEADIVYM 1043

QY 442 QOADYAVEKGILDK 456
||| : : : : : : : : : : : : : : : : :
DB 1044 AVADKAAEAGVTSEQ 1058

RESULT 4
A26655 myosin heavy chain [similarity] - slime mold (Dictyostelium discoideum)
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Dictyostelium discoideum
C>Date: 05-Oct-1988 #sequence revision 05-Oct-1988 #text_change 19-Apr-2002
C:Accession: A26655, A24728, S00250
R:Warwick, H.M.; De Lozanne, A.; Levinwand, L.A.; Spudich, J.A.

Proc. Natl. Acad. Sci. U.S.A. 83, 9433-9437, 1986
 A:Title: Conserved protein domains in a myosin heavy chain gene from Dictyostelium discoideum
 A:Reference number: A26655; MUID:87092266; PMID:3540939
 A:Accession: A26655
 A:Molecule type: DNA
 A:Residues: 1-2116 <WAR>
 A:Cross-references: GB:M14628; GB:M11938; NID:g167834; PIDN:AAA33227.1; PID:g167835
 R:DeLozanne, A.; Lewis, M.; Spudich, J.A.; Levinwand, L.A.
 Proc. Natl. Acad. Sci. U.S.A. 82, 6807-6810, 1985
 A:Reference number: A24728; MUID:86016788; PMID:3901008
 A:Accession: A24728
 A:Molecule type: mRNA
 A:Residues: 2035-2116
 R:Wagle, G.; Noegel, A.; Scheel, J.; Gerisch, G.
 FEBS Lett. 227, 71-75, 1988
 A:Title: Phosphorylation of threonine residues on cloned fragments of the Dictyostelium A:Reference number: S00250; MUID:88112226; PMID:2828113
 A:Accession: S00250
 A:Status: nucleic acid sequence not shown
 A:Molecule type: DNA
 A:Residues: 1734-1893 <WAG>
 C:Comment: The rod domain is highly periodic, containing a pattern of 7-residue repeats
 C:Superfamily: myosin heavy chain; myosin motor domain homology
 C:Keywords: actin binding; ATP; coiled coil; hydrolase; nucleotide binding; P-loop; phosphatase
 F:1-818/Domain: globular head <HED>
 F:89-747/Domain: myosin motor domain homology <MMOT>
 F:179-186/Region: nucleotide-binding motif A (P-loop)
 F:819-2116/Domain: alpha-helical rod <ROD>

Query Match 7.8%; Score 186.5; DB 2; Length 2116;
 Best Local Similarity 21.6%; Pred. No. 0.29;
 Matches 108; Conservative 86; Mismatches 205; Indels 101; Gaps 17;
 Db 26 SODKS--LHNELSAERPLNEQIAEAEEDKIKTYPPEKPGQSNYSFVDNLNLKAI TE 83
 1023 TKDSKSLRQKKLEELKQVEALAEATAKLAQEAANKLQGEYTEL-NEKFNSEVTA 1081
 84 KEIEKERQSISS--PLDNKLNVEDVDSITKRNKLIIDYDSTKSGLDHKKFDDPGLHQL 141
 1082 RSNVEKSKTLESQVAVNNELDEE---KKNR---DALEKKKALDAMLEEMKQDLEST 1134
 142 DGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGLITESQAHTLEDEVAEVLQKI 201
 1135 GGEKKSLLYD-----KVKQESDMEALRNQISLQSTIAKLEKIKTLGEGVARLQGE 1188
 202 SKE--ANNYEEDPNKPTSWTENQAKIPEKVT-----PMAAIDQGLAKGEN 245
 1189 AEQLAKSNVEKQKKVVELDLEKSAQLAEFTAQAQALDKLKKLEQELSEVQTOLSEANN 1248
 246 D-----ETVSNLTITLNGLERKTKTYSEDN-----FEELQYFPNFYALLKSIDSE 290
 1249 KVNVSSTNKHLETSFNNLKLELAEOAKQALEKKRIGLESELKH-----VNEQLEEE 1302
 291 KEAEK-----ETLITIMKTLDIFVMMVKGTITSPBEGSVLE----- 329
 1303 KKQESNEKRVLDLEKEVSELQDTEEVASKAVTEAKNKKESELDEIKRQYADVSSR 1362
 330 --NIDEMIALQTKN-KLEKNATDNISKLPAPSEKSHETD-----STKEAAKMEKEY 380
 1363 DKSVEQLTKLQAKNEELNATAEAGQDRAERSKKAEFLQEEAVKNLEETAKKVAE 1422
 381 GSLKSDTKDDNSNPGKTDPEKGTAEYLAIRKNIEWLKHKHKKNGKEDYDLGKMRDFI 440
 1423 KAMKKAETDYRSTKSELDDAKNVSEQVQIKRLNEE-----LSLSRVL 1467
 441 NKQADAYVEKGIIDKEAE 460
 1468 -EADERCNSAIKAKKTAES 1486

RESULT 5
 B71609
 hypothetical protein PFB0680w - malaria parasite (Plasmodium falciparum)

C:Species: Plasmodium falciparum
 C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
 C:Accession: B71609
 R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
 ; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.;
 Science 282, 1126-1132, 1998
 A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
 A:Reference number: A71600; MUID:99021743; PMID:9804551
 A:Accession: B71609
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-665 <GAR>
 A:Cross-references: GB:AE001410; GB:AE001362; NID:g3845245; PIDN:AAC71925.1; PID:g3845246
 A:Experimental source: clone 3D7
 C:Genetics:
 A:Gene: PFB0680w

Query Match 7.7%; Score 184; DB 2; Length 665;
 Best Local Similarity 22.0%; Pred. No. 0.089;
 Matches 104; Conservative 76; Mismatches 182; Indels 110; Gaps 17;
 Qy 25 GSODKSLHNELSAERPLNEQIAEAEEDKIKTYPPEKPGQSNYSFVDNLNLKAI TEK 84
 Db 160 GKQDISNSAE--NKQDVKGVELEKKEEKISDDHKVEENKKS-----DD 205
 Qy 85 KEIEKERQSISSPLDNKLN---VEDVDSITKRNKLIIDYDSTKSGLDHKKFDDPGLHQL 140
 Db 206 HKVEENKKSDDHKVEENKSDDHKIEVKEVEHEDEDEDEKKEKSENKKNKDKEND 265
 Qy 141 LDGTPLTAEDIVHKIAARIYEENDRAVDFKIVSKLNLGLITESQAHTLEDEVAEVLQKL 200
 Db 266 EDNDEISDEDDVDEEDKNENDIDDDK-----KETDKTHLEEEENEIIEKE 314
 Qy 201 IS-KEANNYEEDPNKPTSW--TENQAKIPEKVTPEMAAIDQGLAKGENDETIVSNLTITLNG 258
 Db 315 FSKKNGKKNKDKTEKSKDIEK-----EKSKDIEKES-----EKSKEKEKS----- 356
 Qy 259 LERRTKTYSBDFEELQYFPNFYALLKSIDSEKAEKETLITIMKTLDIFVMMVKYGT 318
 Db 357 -KQKEKQKQKE-----KSKDIEK-KEKQDIEKESK-DTAKEKEKDKD 402
 Qy 319 ISPEGVSVLENLDEMIALQTKNKLKQATNTNISKLPAPSEKSH-----E 364
 Db 403 IEKES-----KMEKLNKQNDKEDKDDNEKK-----KNDKQDIDHDDNDENDMBEIE 451
 Qy 365 ETSTKEAAKME-----KEYGSLKDSPTK--DNNSNPGKTDPEKGTAEYLAEL 412
 Db 452 ENDEDEDEDMENKKGKNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGN 511
 Qy 413 RKNIEMWLKHKHKKNGKEDYDLGKMRDF-----INQADAYVEKGIID 454
 Db 512 NENNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGNNGN 563

RESULT 6
 S67593
 transport protein USO1 - yeast (Saccharomyces cerevisiae)
 N:Alternate names: protein D2552; protein YDL058w
 C:Species: Saccharomyces cerevisiae
 C:Date: 12-Jul-1996 #sequence_revision 12-Jul-1996 #text_change 21-Jul-2000
 C:Accession: S67593; A38455; S30782
 R:Bloeker, H.; Brandt, P.
 submitted to the Protein Sequence Database, July 1996
 A:Reference number: S67587
 A:Accession: S67593
 A:Molecule type: DNA
 A:Residues: 1-1790 <BLO>
 A:Cross-references: EMBL:274106; NID:g1431058; PID:e253003; PID:g1431059; MIPS:YDL058w
 A:Experimental source: strain S288C
 R:Nakajima, H.; Hirata, A.; Ogawa, Y.; Yonehara, T.; Yoda, K.; Yamasaki, M.
 J. Cell Biol. 113, 245-260, 1991
 A:Title: A cytoskeleton-related gene, USO1, is required for intracellular protein transport
 A:Reference number: A38455; MUID:91185402; PMID:2010462

A;Accession: A38455

A;Molecule type: DNA

A;Residues: 1-389, 'TA', 392-724, 'S', 726-1790 <NAK>

A;Cross-references: GB:X54378; NID:94777; PIDN:CAA38253.1; PID:94778

A;Note: the authors translated the codon ACT for residue 768 as Ile

R;Hottel, M.K.; Herman, D.J.; Bendel, C.M.; McClellan, M.; Tao, N.; Kendrick, K.E.

submitted to the EMBL Data Library, February 1993

A;Description: An integrin analogue in Saccharomyces cerevisiae.

A;Reference number: S30782

A;Accession: S30782

A;Molecule type: DNA

A;Residues: 71-846, 'E', 848-923, 'K', 925-1252, 'I', 1254-1318, 'V', 1320-1460, 'S', 1462-1580, 'S'

A;Cross-references: EMBL:L03188

C;Genetics:

A;Gene: SGD:USO1; INT1

A;Cross-references: SGD:S0002216; MIPS:YDL058w

A;Map position: 4L

C;Keywords: coiled coil; transmembrane protein

F;326-342/Domain: transmembrane #status predicted <TM1>

F;394-410/Domain: transmembrane #status predicted <TM2>

F;617-633/Domain: transmembrane #status predicted <TM3>

Query Match 7.6%; Score 193.5; DB 2; Length 1790;

Best Local Similarity 20.3%; Pred. No. 0.33;

Matches 120; Conservative 91; Mismatches 200; Indels 179; Gaps 23;

Qy 29 KSLHN--RELSAERPLNEQIAEAEEDKIKTYPPENKPGQSNYSFVDMNLNLLKAI---T 82

Db 965 KSLANNYKMQAE--NESLIKAVEE-----SKNESSIQLSNLQNKIDMSQ 1008

Qy 83 EKEIKERQSIKSPDLNKLNVEDVSTNKRKI-----DDYSTKSGLDHFKPD--- 133

Db 1009 EKENFOIERSIEKNIQKKTISDLEQTEIEIISKSDSDEYESQISLKEKLETATT 1068

Qy 134 -DPGLHOLDGTPLEADIVHKIARIYEENDRAVDFKIVSKLNLGLITESQAH----- 187

Db 1069 ANDENVNKISLTKTRBELEAEALAA--YXNKLNELETKTSEKALKEVNEHLEKEEK 1126

Qy 188 -TLEDEVAELQKLISKEAN-----NVEED-PNKPTSWTNOAGKIPEK 229

Db 1127 IOLEKEATETKQQLNSLRANLESLEKEHDLAAQLKYEEQIANKERYNE-EISQLNDE 1185

Qy 230 VTPMAALQDGLAKENDETIVNTLTNGLERRT-----KTYSEDN-- 270

Db 1186 ITSTQENESIKK-KNDELEGEVAMKSTSEQSNKKSEIDALNLQIKELKKNETNEA 1244

Qy 271 -----PELOVFPNF-----YALLKSIDSEKEAK 294

Db 1245 SLLESIKVSESETVKIKELQECNFKKEVSELEDKLSKASEDKNSKYLELQK-ESEKIKE 1303

Qy 295 EKETLITIMKTLDIDFVQMVK----- 315

Db 1304 ELDAKTTTELKIQLEKITNLKAKEKSELSRLKKTSSERKNABEQLEKLNKIQIKNQ 1363

Qy 316 -----YGTISPE--EGVSYLENDEMIALQTKNLEKNAEDNTSKLPPAPSEK 361

Db 1364 AFPEKRLNKGSTITQYSEKINTLE--DELIRLQNELKAKEDINTRSELEKVSLS 1421

Qy 362 SHEETDSTKEAAKMEKEYGSLKDS-TKDDNSNPGKTDPE-----KKGTEAVLEAIRK 414

Db 1422 NDLELLEKQNTIKSLQDEILSYKDKITRNDEKLLSIERNKRDLESLEKQIRAAQESKAK 1481

Qy 415 NIEWLKHKDKGNKEDYLSKMRDFINKQADAYVEKGILD-KEEAEAIKR 463

Db 1482 VERGLKLEESSKEAELEKSEMM-KLESTIESNETELKSSMETIRK 1530

RESULT 7

A56577

A;Title: microtubule-associated protein MAP 1B - rat (fragment)

C;Species: Rattus norvegicus (Norway rat)

C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 16-Feb-1997

C;Accession: A56577

R;Zauner, W.; Kratz, J.; Staunton, J.; Feick, P.; Wiche, G.

Eur. J. Cell Biol. 57, 66-74, 1992

A;Title: Identification of two distinct microtubule binding domains on recombinant rat M

A;Reference number: A56577; MUID:92347374; PMID:1639092

A;Accession: A56577

A;Status: preliminary

A;Molecule type: mRNA

A;Residues: 1-2364 <ZAU>

A;Cross-references: GB:X60550

A;Experimental source: brain

A;Note: nucleotide sequence not given; conceptual translation not complete

C;Superfamily: microtubule-associated protein MAP1B

Query Match 7.5%; Score 180.5; DB 2; Length 2364;

Best Local Similarity 22.7%; Pred No. 0.65;

Matches 120; Conservative 77; Mismatches 200; Indels 131; Gaps 22;

Qy 22 KPGSQDKSLHNRELSAERPLNEQIAEAEEDKIK-KTYPPENKPGQSNYSFVDMNLNLLKA 80

Db 454 KPLSSKSVRKESKEAPEATKASQVEKTPKVESKEKIVKDKPGK-----VESKPS 505

Qy 81 ITEKEKTEKERQSIKSPDLNKLNVEDVSTNKRKLIDYDSTKSGLDHFKFQDDP--DGL 138

Db 506 VTEKEVPSKEQSPVKAFAEAAATESKPKVTKQKVKKEIKTKP--BEKKEEKPKKEVA 563

Qy 139 HQLDGTPLTARDIVHKIAARIYEENDRAVDFKIVSKLNLGLITESQAH--LEDEVAEVL 197

Db 564 KKEDTPLPKDEKPKK-----BEAKKEIKKEKEKELKKEVKETPLKDAKKEV- 615

Qy 198 QKLISKEANNVEEDPNK-----PTSWTENQAG-----KIPEKVTMAA- 235

Db 616 KKDEKKEVKEKEPKKEIKKIDIKKSTPLSDTKKPAALPKPVAKKEEPTKKEPIAAG 675

Qy 236 -IQD-GLAK-----GENDETVSNTL-----TLTNGLERRTK 264

Db 676 KLKDKGKVKYKKGKTEAAATAVGTAATAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 735

Qy 265 TYSNDFNEELQYFPNFVALLKSIDSEKEAK-----KETLITIMKTLDIDFVQMVV 314

Db 736 DFELKAAEIDVAKIDKQLELIEDEKLEKTEGEAVYIQETEVS----- 782

Qy 315 KYGTISPEEGVSYLENDEMIALQTKNKL---EKNATDNISKL-----FPAPSKSHEE 365

Db 783 KGSASEPDGEGITTEGE--CEQTPPELEPVKQGVDDIEKFEDEGAGFESSEAGDYE 840

Qy 366 TDSTKEAAKMEKEYGSLKOSTKDDNSNPGKTDPEKQTEA--YLEAIRKNI----- 416

Db 841 EKAETEABEEDPDG---EDNVSGSKHSPTEDEETIAKAEADVHIKEKRESVASGDDRA 897

Qy 417 -----EWLKKHDKGNKEDYLSKMRDFINKQADAYVEKGILDKEEAE 459

Db 898 EEDMDALEKGEAQSEEEGEE-----DKAEDAREDEHDPDKTEAE 940

RESULT 8

A45605

A;Title: mature-parasite-infected erythrocyte surface antigen MESA - malaria parasite (Plasmodium

C;Species: Plasmodium falciparum

C;Date: 22-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 08-Dec-2000

C;Accession: A45605; A54517

R;Coppel, R.L.

Mol. Biochem. Parasitol. 50, 335-347, 1992

A;Title: Repeat structures in a Plasmodium falciparum protein (MESA) that binds human ery

A;Reference number: A45605; MUID:92158014; PMID:1741020

A;Accession: A45605

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-1526 <COP>

A;Experimental source: Papua New Guinean isolate FC27

A;Note: sequence extracted from NCBI backbone (NCBIN:83648, NCBI:83656)

R;Coppel, R.L.; Culvenor, J.G.; Bianco, A.E.; Crewther, P.E.; Stahl, H.D.; Brown, G.V.; /

Mol. Biochem. Parasitol. 20, 265-277, 1986

A;Title: Variable antigen associated with the surface of erythrocytes infected with matu


```

Db 206 --KNNKMEINDKLNK-IKSPEDIEKLFNENKKNLYEKFINKL-----EKKRALEL 257
QY 170 KIVS-KLNLGLITISQAH-----TLEDEVAEVLQKLISKEANNYEEDPNK 214
Db 258 KNOELKILEYDLNTVVEARETLNRHKBREYKYSIVDBIRKIESRL--RELKSHVEDYVK 315
QY 215 PLSWTENQAGKIPEKVTWMAAIOGLAGENDETWSNTLTITNGLERRTKTYTYSNFEEL 274
Db 316 LTKQLEIIGDI-EK-----LKEFINKSKYRDDIDNLDITLNLKI-----KDEIERV 360
QY 275 QYFPNFYALLKSIDSEKAKK-----ETLITIMKTLIDFVMMVKYGTISPE 322
Db 361 ETIKDLEELKNLNEIEIKYKRICBECKEYVEKYLEBEKAVEYNKLTLEYITL--- 417
QY 323 EGVSYLENDEMIALQTNKLEKNATD--NISKLPAPS-----EKSHEETDSTK-- 370
Db 418 -----LQEKKSIEKNINDLETRINKLLEBETKNIDIESIENSLKETEKKV 463
QY 371 -----EEAAKMEKEVGLKOSTK-----DONSPPGK-----TDBPK-----GKT 405
Db 464 LENLQKEKIELNKLGEINSEIKLKKILDELKEVEGKCPICKTPIDENKKNELINQHT 523
QY 406 E--AYLE-----AIRKNIEWLKKG-----DKKGNKEDYDL----- 433
Db 524 QLNKKYTELEBINKKIREIKEDIKLEKKEIDKBEENLTKLTLYLEKQSQIEELEKLKNY 583
QY 434 SKMRDFINKQADAYVEKG-----ILD--KEBAEAIKRIYS 466
Db 584 KEQJDEINKKISNVINGKPVDEILEIDIKSLQNKFRNFYN 623

RESULT 11
B71603
RESA-H3 antigen PFB0915w - malaria parasite (Plasmodium falciparum)
C:Species: Plasmodium falciparum
C:Date: 13-Nov-1998 #sequence_revision 13-Nov-1998 #text_change 21-Jul-2000
C:Accession: B71603
R:Gardner, M.J.; Tettelin, H.; Carucci, D.J.; Cummings, L.M.; Aravind, L.; Koonin, E.V.;
; Pertea, M.; Salzberg, S.; Zhou, L.; Sutton, G.G.; Clayton, R.; White, O.; Smith, H.O.
Science 282, 1126-1132, 1998
A:Title: Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.
A:Reference number: A71600; MUID:99021743; PMID:9804551
A:Accession: B71603
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-1558 <GAR>
A:Cross-references: GB:AE001424; GB:AE001362; NID:G3845307; PIDN:AACT1972.1; PID:G384530
A:Experimental source: clone 307
C:Genetics:
A:Gene: PFB0915w

```

```

Query Match 7.5%; Score 179; DB 2; Length 1558;
Best Local Similarity 21.5%; Pred. No. 0.45;
Matches 118; Conservative 96; Mismatches 197; Indels 138; Gaps 25;

QY 27 QDSKLNRLSAB--RPLNEQIAEAEDEKIKKTYPPENKPGQSN-----YSFVDNL-- 75
Db 725 EENAVESNENVAENLEKLNATVENTVLDKVEETVEISGESLENEMDKAFFSEIFDNVKG 784
QY 76 ---NLK-----ALTEKEKIEKQSRSSPLDNKLN-----VEDVDSTK-- 112
Db 785 IQENLITGFRSTETISIVIQSEKVDL--NENVVSSILDNIENKKEGLNKLNIESTEGV 843
QY 113 -----NRKLDDYD-----STKSGLDHKQDDPDGLHQLDGPPLTAEDI- 151
Db 844 QETVTEHVEQNVVDVDPAMKDQPLGILNEAGGLKEMEFNLED-VFKSESVDIVIEIK 902
QY 152 -----VHKIAARIYEENDRAVDFKIVSKLNLGLITISQAHTEDEVAEVLQKLISKE 204
Db 903 DEPVQKEVSKETVTSIIEENIVD-----VLEEKEDLTDKMDAVERISIBIS 951
QY 205 ANNYEEDPNKPTSWTENQAGKIPEKVTWMAAIOGLAGENDETWSNTLTITNGLERITK 264

```

```

Db 952 SDSKEE-----TES-----TKQEKQVSLVVEEQDNDMDSEVKVLKLNMEELMK 999
QY 265 TYSEDN-----FEELQYFPNFYA-LLKSIDSEKAK-----EKETLITIMK 304
Db 1000 DAVEINDITSKLIEETQELNEVEADLIKMEKLEKALSSEDSKEIIDAKDITLKVIE 1059
QY 305 -----TLIDFVMMVKYGTISPE--EGVSYLENDEMIALQTK--NKLKKNATDNISKL 354
Db 1060 EEHDITITLDEV---VELKQVBEEDKIEKVDLQLEEDILKEVKEIKELSEBILEDYKEL 1116
QY 355 FPAPSSKSEHEEDTSTKEBAAKMEKEVGLKOSTKDNNSPPGKTDPEPKCTAYLEAIRK 414
Db 1117 KTIETDILKEKKEIKDHPEKFEAEETKQLEADILKEVSVLESVEEEKKLEBEVHE-LKE 1175
QY 415 NIWLKHKDK--KGNKEDYDLKMRDFINKQADAY---VEKGILDK-----EBA 458
Db 1176 EVEHIISGDAHIKGLEED-DLEEVDDLKGSILDKMLKGMELGDMDKESLVEDVTAKLGERV 1234
QY 459 EAIKRIYS 467
Db 1235 ESLKQVLSS 1243

RESULT 12
T28677
rhoptry protein - plasmodium yoelii
C:Species: Plasmodium yoelii
C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 01-Dec-2000
C:Accession: T28677; C45521
R:Keen, J.; Sinha, K.; Brown, K.; Holder, A.
Mol. Biochem. Parasitol. 65, 171-177, 1994
A:Title: A gene coding for a high molecular mass rhoptry protein of Plasmodium yoelii.
A:Reference number: Z20508; MUID:95021522; PMID:7935623
A:Accession: T28677
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-2269 <KEE>
A:Cross-references: EMBL:L27838; NID:9457145; PID:9457146; PIDN:AAA21304.1
R:Keen, J.; Holder, A.; Playfair, J.; Lockyer, M.; Lewis, A.
Mol. Biochem. Parasitol. 42, 241-246, 1990
A:Title: Identification of the gene for a Plasmodium yoelii rhoptry protein. Multiple co
A:Reference number: A45521; MUID:91101660; PMID:2270106
A:Accession: C45521
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 2131-2269 <KE2>
A:Cross-references: GB:M34283

Query Match 7.5%; Score 179; DB 2; Length 2269;
Best Local Similarity 21.1%; Pred. No. 0.73; Mismatches 85; Indels 144; Gaps 23;
Matches 113; Conservative 85;

QY 13 LVLPQAPFKPGSQDLSLHNRLSAERPLNEQIAEAEDEKIKKTYPPENKPGQSNYSFV 72
Db 7 LILPLNQFKGLNESMIKLNKSGILRKYTISNQI---KNKLVNSTYPEGREG-----FT 57
QY 73 DNALN-----LKAITEKEK-----TEKERQSTRSSPLDN-----KLN- 104
Db 58 SSLELAKSWKTKLETITELTKNEETVRLEKEIRELFKYLDEAEARKYLEGLKLELNK 117
QY 105 -----VEDVDST-----KNRKLIDD-----YDSTKSLGDH 129
Db 118 KIKDIIAKIEVKNVTVELKKEIKKNAYIDELANQSPYKVTGYENKNTYNTIKSFQD 177
QY 130 KFQDDPGLHQLDGTPLTAEDIVHKIARIYBENDRAVDFKIVSKLNLGLIT-ESQAH 188
Db 178 IYEGDIDTFYN-ELSSIVKEDPIDDEDKTKLENLRSKIDNVYDKIQKMEIETVKSHLN 236
QY 189 LE-----DEVAEVLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVT 231
Db 237 IETNNKLPVILIKKIYDIYSKELNKLMLDPFNKEKLSNKLISDY-DKKRQQLSEYKS 295

```

```
QY 232 PMAAIOGLAKENDET-VSNLTITLNGLERKTKYSEDNFELOYPFNFVALLKSIDSE 290
Db 236 KMLEIRNH-NSQTNVDNTE-----EAAKQYDKSN-EHMTTITPTNEDEISIISE 345
QY 291 KEAKKETLITIMKTLDIPVKMMVKGITSPREGVSYLENDEMIALOTKMKLEKNATDN 350
Db 346 VKTM-KDELLSKVNTYIDNK---KYKTVNSEHQSOTFELTDKIKAEVSDKELKK----- 396
QY 351 ISKLPAPAPESHEETS-TKEAAAMEKEYGSLKSDTKDONSPPGKTDPPKGTAYL 409
Db 397 -----CEQSFNDKSLINETKNSIEKEY-----QNINTLKKVDE-----YI 432
QY 410 EAIRKNIEWLKHKDKGNKEDYDLKMRDPIFKQADAVVEKGIIDKEEABAIRI 464
Db 433 KVCKSTKESITKFSK-----QTILKMLNQNITKVTNSIDSKYIERFEQI 480

RESULT 13
A43336
Query Match 7.4%; Score 178.5; DB 2; Length 1392;
Best Local Similarity 22.4%; Pred. No. 0.42;
Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C:Accession: A43336
R:Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
Cell 70, 887-900, 1992
A:Title: CLIP-170 links endocytic vesicles to microtubules.
A:Reference number: A43336; MUID:92405160; PMID:1356075
A:Accession: A43336
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1392 <PIE>
A:Cross-references: GB:M97501; NID:g180621; PIDN:AAA35693.1; PID:g180622

Query Match 7.4%; Score 178.5; DB 2; Length 1392;
Best Local Similarity 22.4%; Pred. No. 0.42;
Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;
C:Species: Homo sapiens (man)
C>Date: 30-Sep-1993 #sequence_revision 30-Sep-1993 #text_change 05-Nov-1999
C:Accession: A43336
R:Pierre, P.; Scheel, J.; Rickard, J.E.; Kreis, T.E.
Cell 70, 887-900, 1992
A:Title: CLIP-170 links endocytic vesicles to microtubules.
A:Reference number: A43336; MUID:92405160; PMID:1356075
A:Accession: A43336
A>Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-1392 <PIE>
A:Cross-references: GB:M97501; NID:g180621; PIDN:AAA35693.1; PID:g180622

QY 34 RELSAE-----RPLNEQIAEAEEDKIKTYPENKPGQSNYSFVDNL-----NLLKA 80
Db 734 RKASSEGKSEMKLQQL-EAAEKQIKHLEIEKNAESSKASSITRELQGRELKLTNLQEN 792
QY 81 ITE-----KEKIEKERQ-----SIRSSPLD--NKL-----NVEDVDSTKN 113
Db 793 LSEVSQVKETLEKELQILKEKFAEAEAVSVQSMQETVNLKHQKEQFNMSSDLEKL 852
QY 114 RKLIDYDSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVS 173
Db 853 RENLAD-----MEAKFREKDEREEQL-----IKAKEKLEND-----IA 885
QY 174 KLLNLGLITESQAHTLEDEV-----AEVLQKLISKEANNYE-----EDPNKPTSWTEN 221
Db 886 EIMKSGDSSQLTKMDELKRDVEELQKLTKANENASFLOKSTIEDTVKAEQSQ 945
QY 222 QAGKIPKVTMAAIOGLAKENDETNSNTLTNGLERKTKYSEDNFELOYPFNFY 281
Db 946 EAAKKEE-----EKKELERKLSL-----LEKKMET-SHNQCOELK----- 980
QY 282 ALLKSIDSEKAKEKETLITIMKTLDIPVKMMVKGITSPREGVSYLENDEMIALQTKN 341
Db 981 ARYERATSETKTHEEILQNLQKLTLDTEDEKLG-----ARENSGLLQELBELRQAESA 1036
QY 342 KLEKNATDNISKLPAPSEKSH--EETDSTKEAAKMEKEYGSLKSDTKDONSPPGKTD 399
Db 1037 KAAQTAEADAMQIMEQMTKEKTETTLASLEDTKQTNKLNQELDTLKENNL-KNVEELNKS 1095
QY 400 EPKGTKEAYLEAIRKNIEWLKSH-----DKGNKEDYDLKMRDPIFKQADAVV 448
Db 1096 ELLTVENQKMEEFREKIEITLKQAAQKSQLSALQENVKLAEBELGRSDRDEVTSHOKLEE 1155
QY 449 EKGILDKKEEAAIKR 463
Db 1156 ERSVLNNQLEMKKR 1170
```

RESULT 14

S22695

restin - human

C:Species: Homo sapiens (man)

C>Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 05-Nov-1999

C:Accession: S22695; S19853

R:Bilbe, G.; Delabie, J.; Brueggen, J.; Richener, H.; Asselbergs, F.A.M.; Cerletti, N.;

EMBO J. 11, 2103-2113, 1992

A:Title: Restin: a novel intermediate filament-associated protein highly expressed in the

A:Reference number: S22695; MUID:92289675; PMID:1600942

A:Accession: S22695

A:Molecule type: mRNA

A:Residues: 1-1427 <BIL>

A:Cross-references: EMBL:X64838; NID:g35998; PIDN:CAA46050.1; PID:g35999

C:Keywords: cytoskeleton

Query Match 7.4%; Score 178.5; DB 2; Length 1427;

Best Local Similarity 22.4%; Pred. No. 0.43;

Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;

QY 34 RELSAE-----RPLNEQIAEAEEDKIKTYPENKPGQSNYSFVDNL-----NLLKA 80

Db 769 RKASSEGKSEMKLQQL-EAAEKQIKHLEIEKNAESSKASSITRELQGRELKLTNLQEN 827

QY 81 ITE-----KEKIEKERQ-----SIRSSPLD--NKL-----NVEDVDSTKN 113

Db 828 LSEVSQVKETLEKELQILKEKFAEAEAVSVQSMQETVNLKHQKEQFNMSSDLEKL 887

QY 114 RKLIDYDSTKSGLDHKFQDDPGLHQLDGTPLTAEDIVHKIAARIYEENDRAVFDKIVS 173

Db 888 RENLAD-----MEAKFREKDEREEQL-----IKAKEKLEND-----IA 920

QY 174 KLLNLGLITESQAHTLEDEV-----AEVLQKLISKEANNYE-----EDPNKPTSWTEN 221

Db 921 EIMKSGDSSQLTKMDELKRDVEELQKLTKANENASFLOKSTIEDTVKAEQSQ 980

QY 222 QAGKIPKVTMAAIOGLAKENDETNSNTLTNGLERKTKYSEDNFELOYPFNFY 281

Db 981 EAAKKEE-----EKKELERKLSL-----LEKKMET-SHNQCOELK----- 1015

QY 282 ALLKSIDSEKAKEKETLITIMKTLDIPVKMMVKGITSPREGVSYLENDEMIALQTKN 341

Db 1016 ARYERATSETKTHEEILQNLQKLTLDTEDEKLG-----ARENSGLLQELBELRQAEDA 1071

QY 342 KLEKNATDNISKLPAPSEKSH--EETDSTKEAAKMEKEYGSLKSDTKDONSPPGKTD 399

Db 1072 KAAQTAEADAMQIMEQMTKEKTETTLASLEDTKQTNKLNQELDTLKENNL-KNVEELNKS 1130

QY 400 EPKGTKEAYLEAIRKNIEWLKSH-----DKGNKEDYDLKMRDPIFKQADAVV 448

Db 1131 ELLTVENQKMEEFREKIEITLKQAAQKSQLSALQENVKLAEBELGRSDRDEVTSHOKLEE 1190

QY 449 EKGILDKKEEAAIKR 463

Db 1191 ERSVLNNQLEMKKR 1205

RESULT 15

D96796

probable heat shock protein, 53413-59028 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001

C:Accession: D96796

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chen, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.;

anssen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khavkin, E.; Kim, C.C.

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96796
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1871 <STO>
A:Cross-references: GB:AE005173; NID:G6143906; PIDN:AAF04452.1; GSPDB:GN00141
C:Genetics:
A:Gene: F28016.15
A:Map position: 1

```
Query Match          7.4%; Score 177; DB 2; Length 1871;
Best Local Similarity 20.7%; Pred. No. 0.71;
Matches 116; Conservative 85; Mismatches 212; Indels 148; Gaps 22;

QY 12 VLVLPQAFKPP--GGSODKSLHNLRELSAERPLNEQTAEAEEDKIKTYPPENKPGOSNY 69
Db 1052 VLVVEETYPKDTGGDHDNDHKEEQENKVIKAEINTEEDSFKKVEIEKQ-----1105

QY 70 SFVDNLNLLKAITKEKIEKERQIRSGSPLDNKLNVEDVDSTKN---RKLDDYDSTKSG 126
Db 1106 ---DHGELKRSVMQAKRQETEEKD-KTFAMEKNETVERRKQTKDGSGLKREGEDPELGG 1161

QY 127 LDHKFQDDPGLHQLDGTPLT--AEDIVHKTAARIYEENDRAVED----KIVSKLLNLGL 180
Db 1162 --HERRGEEDRIEELVETIEISDHKEKVKKKOEDYILRSQDTGKVDLGERERRRSKQRKIHK 1219

QY 181 ITESQAHTEDEVAEVLQKLSKEAN-----NYEEDPNKPTSWTENQAGKIPEKVTP- 232
Db 1220 SVEDEIGDQDEDAEAAAVSVRNENGSSRRKVQTIEESEK-----HKEQNKIPETSNPE 1274

QY 233 -----MAAIQDGLAKGEN-----DETVSNTLTLTNGLERRTKT 265
Db 1275 VNEEDEERVVEKETKEVEAHVQELEGTENCKDDGGRREERGKQGMTAENMLRQRFKT 1334

QY 266 YSEDNPELQYFPNFPYALLKSID--SEKAKEKETLITIMKTLDIVKMMVKYGTI---- 319
Db 1335 KSDD-----GIVRKIQETKEEPDEKKS-----QESSSHVVVKLVAEEDGSLRNL 1378

QY 320 ---SPEEGSVYLENDWMIALQTKNLEKNATD--NISKLFPAPSEKSHETDST----- 369
Db 1379 EFSEKSTVSKMLKLDSEKKEEHKIRKPTERSNAPVIEKQGNKKNKAAEENQDKIDRR 1438

QY 370 -KEEAAKMEKEYGSLKD-----STKDDNSNPGGKTD-- 399
Db 1439 GKQOEIKGQBPYGLVRNGEHDKITEVHRGEEKGTAENVSSTKIQTQKDELEKPKPSEIS 1498

QY 400 -----EPKGTAEYLEATRKNIEMWLKHDHKKGNKEDYDLS--KMRDFIN 441
Db 1499 ENNHIEFMDSSQSDIEEGSDQAEKYAKQNKIQEVMNDEK--KEEYHISERVNEMA 1556

QY 442 K---QADAYVEKGILDKKEEAE 459
Db 1557 KRILQVESKANDGSSKKNETE 1577
```

Search completed: January 2, 2004, 11:02:28
Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 10:59:27 ; Search time 17 Seconds
(without alignments)
1294.618 Million cell updates/sec

Title: ABU66614

Perfect score: 2399

Sequence: 1 MGFLGTGTLVLVLPIQAF.....EKGILDKERAEAKRIYSSL 468

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 127863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_41.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2395	99.8	468	1 SG3 HUMAN	Q8wxd2 homo sapien
2	2102.5	87.6	471	1 SG3 MOUSE	P47867 mus musculus
3	2096.5	87.4	471	1 SG3 RAT	P47868 rattus norv
4	191	8.0	2464	1 MAPB MOUSE	P14873 mus musculus
5	186.5	7.8	2116	1 MYS2 DICDI	P08799 dictyosteli
6	183.5	7.6	1790	1 US01 YEAST	P25386 saccharomyc
7	180.5	7.5	2459	1 MAPB RAT	P15205 rattus norv
8	179	7.5	1005	1 RA50 METJA	Q58718 methanococc
9	178.5	7.4	1427	1 REST HUMAN	P30622 homo sapien
10	178	7.4	882	1 RA50 PYRFU	P58301 streptococc
11	172.5	7.2	1164	1 BAG STRAG	P27951 pyrotoconc
12	169	7.0	700	1 TRDN CANPA	P82179 canis famil
13	167.5	7.0	957	1 KPSC HUMAN	O60282 homo sapien
14	166.5	6.9	1875	1 MFP1 YEAST	Q02455 saccharomyc
15	166	6.9	2663	1 CENO HUMAN	Q02224 homo sapien
16	166	6.9	5038	1 PCLO MOUSE	Q9gyx7 mus musculus
17	163.5	6.8	956	1 KPSC MOUSE	P28738 mus musculus
18	163	6.8	1251	1 RPF2 PLAYB	Q00799 plasmodium
19	161.5	6.7	1360	1 CING XENLA	Q9ptd7 xenopus lae
20	161.5	6.7	1969	1 MYS4 CAEEL	P12844 caenorhabdi
21	161	6.7	852	1 RA50 THENA	Q9x1x1 thermotoga
22	159.5	6.6	997	1 SCP1 RAT	Q03410 rattus norv
23	159.5	6.6	1978	1 MYHB CHICK	P10587 gallus gall
24	159	6.6	1208	1 PCP1 SCHPO	Q92351 schizosacch
25	158.5	6.6	1526	1 MYS2 SCHPO	Q9usi6 schizosacch
26	158	6.6	1972	1 MYHB RABIT	P35748 oryctolagus
27	157.5	6.6	1972	1 MYHB HUMAN	P35749 homo sapien
28	157.5	6.6	2230	1 GCG4 HUMAN	Q13439 homo sapien
29	157	6.5	727	1 MFP1 ARATH	Q91w85 arabidopsis
30	157	6.5	944	1 NUF1 YEAST	P32380 saccharomyc
31	157	6.5	2017	1 MYSN DROME	Q99323 drosophila
32	156.5	6.5	1087	1 AKA9 RABIT	Q28628 oryctolagus
33	156.5	6.5	2468	1 MAPB_HUMAN	P46821 homo sapien

RESULT 1

SG3_HUMAN

ID SG3_HUMAN STANDARD; PRT; 468 AA.

AC Q8WXD2; Q96C83; Q96GE8; Q9Y6G7;

DT 28-FEB-2003 (Rel. 41, Created)

DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 15-SEP-2003 (Rel. 42, Last annotation update)

DE Secretogranin III precursor (SgIII).

GN SGC3.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

OX NCBI_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Pituitary;

RA Song H., Peng Y., Huang Q., Dai M., Mao Y., Zhang Q., Mao M., Fu G.,

RA Luo M., Chen J., Hu R.;

RT "Human secretogranin III mRNA, complete cds.";

RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A. AND TISSUE SPECIFICITY.

RX MEDLINE=22093564; PubMed=12098761;

RT Rong Y.P., Liu F., Zeng L.C., Ma W.J., Wei D.Z., Han Z.G.;

RT "Cloning and characterization of a novel human secretory protein:

RL secretogranin III.";

RN Acta Biochim. Biophys. Sin. 34:411-417(2002).

RN [3]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain, and Lung;

RX MEDLINE=22388257; PubMed=12477932;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,

RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.B., Scheetz T.E.,

RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,

RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,

RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,

RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,

RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,

RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,

RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,

RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

RT "Generation and initial analysis of more than 15,000 full-length

RL human and mouse cDNA sequences.";

RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).

CC -1- SUBCELLULAR LOCATION: Secreted. Neuroendocrine and endocrine

CC secretory granules.

CC -1- TISSUE SPECIFICITY: Expressed in brain, heart, kidney, liver and

CC skeletal muscle.

OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
ON NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=95001263; PubMed=7917832;
RA Dopazo A., Lovenberg T.W., Danielson P.E., Ottiger H.-P.,
RA Sutcliffe J.G.;
RT Primary structure of mouse secretogranin III and its absence from
RT deficient mice.;
RL J. Mol. Neurosci. 4:225-233(1993).
RN [2]
RP PRELIMINARY SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90376160; PubMed=2204688;
RA Ottiger H.-P., Battenberg E.F., Teou A.-P., Bloom F.E.,
RA Sutcliffe J.G.;
RT "B1075; a brain- and pituitary-specific mRNA that encodes a novel
RT chromogranin / Secretogranin-like component of intracellular
RT vesicles.";
RL J. Neurosci. 10:3135-3147(1990).
CC -I- SUBCELLULAR LOCATION: Secreted. Neuroendocrine and endocrine
CC secretory granules.
CC -I- TISSUE SPECIFICITY: BRAIN- AND PITUITARY-SPECIFIC.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; U02983; AAA56637.1; --
KW Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 22 POTENTIAL.
FT CHAIN 23 471 SECRETOGRANIN III.
SQ SEQUENCE 471 AA; 53183 MW; 977B3F885F33223 CRC64;
Query Match 87.4%; Score 2096.5; DB 1; Length 471;
Best Local Similarity 87.9%; Pred No. 1.1e-93;
Matches 414; Conservative 18; Mismatches 36; Indels 3; Gaps 1;
QY 1 MGFLGTGWLVLVL---PIQAPFGSGQDSKLNHLSAERPLNEQIAEAEADKIKKT 57
Db 1 MGFLWTGSWILVLVLSGPIQAPFGSGQDSKLNHLSAERPLNEQIAEAEADKIKKT 60
QY 58 YPPENKPGQSNYSFVDNLNLKAI TEKEKIKEROSIRSSPLDNKLVNEDVDSTNKRKLI 117
Db 61 YPSEKPSSESNSFSDNLNLKAI TEKETVBEKAKQSISSPFDNRLNVDADSTNKRKLT 120
QY 118 DDYDSTKSLDHFODDPDGLHOLDGTPLEADIVHKAARIYENDRAVFDKIVSKLN 177
Db 121 DEYDSTKSLDGRKQDDPDGLHOLDGTPLEADIVHKAARIYENDRAGVFDKIVSKLN 180
QY 178 LGLITESQAHTLEDEVAEVLQKLSKEANNYEEDPNKPTSWTENQAGKIPKVPMAAIQ 237
Db 181 LGLITESQAHTLEDEVAEVLQKLSKEANNYEAEPEKPTSTENQDGKIPKVPVAAQ 240
QY 238 DGLAKGENDETVSNLTITNGLERTKYTSBDNFEELYQFNFVALLKSIDSEKEAKE 297
Db 241 DGFTRNDDTVSNLTITLSGLERTNPHRDDFEELYQFNFVALLTIDSEKEAKEKE 300
QY 298 TLITIMKTLIDFKVMVKYGTISPFGVSYLENDEMIALQTKNLEKQATDNTSKLPFA 357
Db 301 TLITIMKTLIDFKVMVKYGTISPFGVSYLENDEMIALQTKNLEKQATDNTSKSKLPFA 360
QY 358 PSEKSHETDSTKEAAKWEKEYSLKDSKDDNSNPGGKTDEPKGTEAYLEAIRKNIE 417
Db 361 PPEKSHETDSTKEAAKWEKEYSLKDSKDDNSNPGGKTDEPKGTEAYLEAIRKNIE 420
QY 418 WLKKGDKGNKEDYDLSKQRDFINQADAYVEKGILDKKEAAIKRIYSSL 468

Db 421 WLKKGDKGNKEDYDLSKQRDFINQADAYVEKGILDKKEAAIKRIYSSL 471
RESULT 4
MAPB MOUSE
ID MAPB MOUSE STANDARD; PRT; 2464 AA.
AC P14873;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Microtubule-associated protein 1B (MAP 1B) (MAP1.2) (MAPI(X))
DE [Contains: MAPI light chain LC1].
GN MAPIB OR MTAPIB OR MTA5.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND DOMAIN.
RC STRAIN=Swiss Webster; TISSUE=Brain;
RX MEDLINE=90094539; PubMed=2480963;
RA Noble M., Lewis S.A., Cowan N.J.;
RT "The microtubule binding domain of microtubule-associated protein
RT MAPIB contains a repeated sequence motif unrelated to that of MAP2
RT and tau.";
RL J. Cell Biol. 109:3367-3376(1989).
CC -I- FUNCTION: THE FUNCTION OF BRAIN MAPS IS ESSENTIALLY UNKNOWN.
CC PHOSPHORYLATED MAPIB MAY PLAY A ROLE IN THE CYTOSKELETAL CHANGES
CC THAT ACCOMPANY NEURITE EXTENSION. POSSIBLY MAPIB BINDS TO AT LEAST
CC TWO TUBULIN SUBUNITS IN THE POLYMER, AND THIS BRIDGING OF SUBUNITS
CC MIGHT BE INVOLVED IN NUCLEATING MICROTUBULE POLYMERIZATION AND IN
CC STABILIZING MICROTUBULES.
CC -I- SUBUNIT: 3 DIFFERENT LIGHT CHAINS, LC1, LC2 AND LC3, CAN ASSOCIATE
CC WITH MAPIA AND MAPIB PROTEINS.
CC -I- DOMAIN: Has a highly basic region with many copies of the sequence
CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
CC responsible for the binding of MAPIB to microtubules.
CC -I- PTM: LC1 IS COEXPRESSED WITH MAPIB. IT IS A POLYPEPTIDE GENERATED
CC FROM MAPIB BY PROTEOLYTIC PROCESSING. IT IS FREE TO ASSOCIATE WITH
CC BOTH MAPIA AND MAPIB. IT INTERACTS WITH THE AMINO-TERMINAL REGION
CC OF MAPIB.
CC -I- SIMILARITY: TO MAPIA.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
CC or send an email to license@isb-sib.ch).
CC
EMBL; X51396; CAA35761.1; --
PIR; S07549; QRMSP1.
MGD; MGI:1306778; Mtapib.
GO; GO:0016358; P:dendrite morphogenesis; IMP.
GO; GO:0001578; P:microtubule bundling; IMP.
InterPro; IPR000102; MAPIB neuraxin.
Pfam; PF00414; MAPIB neuraxin; 10.
DR PROSITE; PS00230; MAPIB_NURAXIN; 7.
KW Microtubules; Repeat; Phosphorylation.
FT CHAIN ? 2464 MAPI LIGHT CHAIN LC1.
FT REPEAT 1874 1890 MAPIB 1.
FT REPEAT 1891 1907 MAPIB 2.
FT REPEAT 1908 1924 MAPIB 3.
FT REPEAT 1925 1941 MAPIB 4.
FT REPEAT 1942 1958 MAPIB 5.
FT REPEAT 1959 1975 MAPIB 6.
FT REPEAT 1993 2009 MAPIB 7.
FT REPEAT 2010 2026 MAPIB 8.
FT REPEAT 2027 2043 MAPIB 9.
FT REPEAT 2044 2060 MAPIB 10.
FT DOMAIN 589 787 LYS-RICH (HIGHLY BASIC, CONTAINS MANY

FT	SQ	SEQUENCE	2464 AA;	270408 MW;	FBD3DD99CFD8DA87 CRC64;	KKEE AND KKEI/V REPEATS).
		Query Match	8.0%;	Score 191;	DB 1;	Length 2464;
		Best Local Similarity	21.3%;	Pred. No. 0.098;		
		Matches 118;	Conservative	85;	Mismatches 186;	Indels 16;
Qy	22	KPGSGQDKSLHNRLSALRPLNQIAA	-----EEDKIKKTYP--PENKPKP			
Db	550	KPVASVRKESKEETPEVTKTSQVKT	PKVESKEVLVKKDKPVTKESKP			
Qy	74	NLNLKAITEKIEKIERQSI	RSSPLDNKLNVEDVDSTKRNKLIDYDSTKK			
Db	601	-----SVTEKYSKKEEQSPVKA	EAQATSKPKVTDKVVKKEIKTK			
Qy	134	DP-----DGLHOLDGTP	L-----TAEDIVHKIAARIYEENDRAVFDKIVSKLL			
Db	653	KPKFVVVKEDKTP	LKKDKPKRFEVKKEIKKEI	KKE-ERKELKKEVKK-		
Qy	186	AHTLEDEVAEVLQKLISK	EANNYEEDPNK-----PTSWTENQAAAG			
Db	704	LKDAKKEVKEKEKVKKE	-----EKEPKKEIKKISKDIKKSTPQSDTKPKS			
Qy	229	-----KVTPLAA--IQD-GLAK	-----GENDETVSNTL-			
Db	760	ESSTKKEPLAAGKL	DKGKVVKVKKSGKTEAAATAVGTAATTA	AAVVAAG		
Qy	254	-----TLTNGLERRTKTYG	EDNFEELQYFPNFVALLKSIDSEKAKE	-----		
Db	820	LEAERSLMSPEDLTKD	PELKAEEIDVAKIKPOLELIEDEKLKETQPG			
Qy	299	LITIMKTLIDFVQWVKYGT	ISPREGVSVLENDEMIALQTKNKL	-----EKN		
Db	880	EVS-----KGSAES	PDEGITTTTEGBGE--CEQTPEELEPVEKQ			
Qy	355	-----FPAPSEKSHETD	STKEEAARWEKYSGLKSDTKDDNSNPGGKTDE			
Db	925	DEGAGPESSETGYDEKA	ETEAESPRED--GEDNAGSGASKHSPTDEDDES			
Qy	410	EAIRKNI-----EWLKK	-----HDKKKKEDY-----D			
Db	984	KEKRESVVGDDRAED	MDVLEKGAEQSBEGBEEDKAEDAREGYPED			
Qy	442	KQADAYVEKILDK	E 456			
Db	1044	AVADKAAEAGVTEEQ	1058			

RESULT 5
MYS2 DICDI
ID MYS2 DICDI STANDARD; PRT; 2116 AA.
AC P08799;
DT 01-NOV-1988 (Rel. 09, Created)
DT 01-OCT-1989 (Rel. 12, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Myosin II heavy chain, non muscle.
GN MHCA.
OS Dictyostelium discoideum (Slime mold).
OC Eukaryota; Mycetozoa; Dictyostellida; Dictyostelium.
OX NCBI_TaxId=44689;
[1]
RN SEQUENCE FROM N.A.
RP MEDLINE=87092266; PubMed=3540939;
RR Warrick H.M., de Lozanne A.; Levinwand L.A., Spudich J.A.;
RA "Conserved protein domains in a myosin heavy chain gene from
RT Dictyostelium discoideum.";
RL Proc. Natl. Acad. Sci. U.S.A. 83:9433-9437(1986).
[2]
PHOSPHORYLATION SITES, AND MUTAGENESIS.
RC STRAIN=AX2;
RX MEDLINE=90853583; PubMed=2387408;
RA Lucke-Vielmeier D., Schleicher M., Grabatin B., Wippler J., Gerisch G.;

RT "Replacement of threonine residues by serine and alanine in a
RL phosphorylatable heavy chain fragment of Dictyostelium myosin II.";
RL PBBS Lett. 269:239-243 (1990).
RN [3]
RP PHOSPHORYLATION SITES.
RX MEDLINE=88112226; PubMed=2828113;
RA Wagle G., Noegel A., Scheel J., Gerisch G.;
RT "Phosphorylation of threonine residues on cloned fragments of the
RL Dictyostelium myosin heavy chain.";
RL PBBS Lett. 227:71-75 (1988).
RN [4]
RX X-RAY CRYSTALLOGRAPHY (2.6 ANGSTROMS) OF 1-762.
RL MEDLINE=95345066; PubMed=7619795;
RA Fisher A.J., Smith C.A., Thoden J.B., Smith R., Sutoh K., Holden H.M.,
RT Rayment I.;
RL "X-ray structures of the myosin motor domain of Dictyostelium
RL Biochemistry 34:8960-8972 (1995).
RN [5]
RX X-RAY CRYSTALLOGRAPHY (2.7 ANGSTROMS) OF 1-762.
RL MEDLINE=95345067; PubMed=7619796;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II)-pyrophosphate complex of the
RL truncated head of Dictyostelium discoideum myosin to 2.7-A
RL resolution.";
RL Biochemistry 34:8973-8981 (1995).
RN [6]
RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RL MEDLINE=96206189; PubMed=8611530;
RA Smith C.A., Rayment I.;
RT "X-ray structure of the magnesium(II).ADP.vanadate complex of the
RL Dictyostelium discoideum myosin motor domain to 1.9-A resolution.";
RL Biochemistry 35:5404-5417 (1996).
RN [7]
RX X-RAY CRYSTALLOGRAPHY (2.1 ANGSTROMS) OF 1-762.
RL MEDLINE=97452580; PubMed=9305951;
RA Gulick A.M., Bauer C.B., Thoden J.B., Rayment I.;
RT "X-ray structures of the MgADP, MgATPgammag, and MgAMPPNP complexes
RL of the Dictyostelium discoideum myosin motor domain.";
RL Biochemistry 36:11619-11628 (1997).
RN [8]
RX X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 1-762.
RL MEDLINE=98070605; PubMed=9405148;
RA Bauer C.B., Kuhlman P.A., Bagshaw C.R., Rayment I.;
RT "X-ray crystal structure and solution fluorescence characterization
RL of Mg.2(3')-O-(N-methylanthraniloyl) nucleotides bound to the
RL Dictyostelium discoideum myosin motor domain.";
RL J. Mol. Biol. 274:394-407 (1997).
CC -1- FUNCTION: MYOSIN IS A PROTEIN THAT BINDS TO ACTIN & HAS ATPASE
CC ACTIVITY THAT IS ACTIVATED BY ACTIN
CC -1- SUBUNIT: MYOSIN II HEAVY CHAIN IS TWO-HEADED. IT SELF-ASSEMBLES
CC INTO FILAMENTS. HEXAMER OF 2 HEAVY CHAIN SUBUNITS (MHC), 2 ALKALI
CC LIGHT CHAIN SUBUNITS (MLC) AND 2 REGULATORY LIGHT CHAIN SUBUNITS
CC (MLC-2).
CC -1- SUBCELLULAR LOCATION: HIGHEST CONCENTRATION IN THE POSTERIOR CELL
CC CORTEX.
CC -1- DOMAIN: EACH MYOSIN HEAVY CHAIN CAN BE SPLIT INTO 1 LIGHT
CC MEROMYOSIN (LMM) AND 1 HEAVY MEROMYOSIN (HMM). IT CAN BE FURTHER
CC SPLIT INTO 2 GLOBULAR SUBFRAGMENTS (S1) AND 1 ROD-SHAPED
CC SUBFRAGMENT (S2).
CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE. SHOWING,
CC CYCLES OF A 28-RESIDUE REPEAT PATTERN COMPOSED OF 4 HEPTAPEPTIDES,
CC CHARACTERISTIC FOR ALPHA-HELICAL COILED COILS.
CC -1- PTM: PHOSPHORYLATION INHIBITS THICK FILAMENT FORMATION AND REDUCES
CC THE ACTIN-ACTIVATED ATPASE ACTIVITY.
CC -1- MISCELLANEOUS: DICTYOSTELIUM MYOSIN II HAS NO K(2)EDTA ATPASE
CC ACTIVITY, PERHAPS CORRELATED WITH THE ABSENCE OF A CYS AT THE SH-1
CC POSITION (688).
CC -1- SIMILARITY: Contains 1 myosin-like globular head domain.
CC -1- SIMILARITY: Contains 1 IQ domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collabora
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation
CC

RESULT 6

USO1_YEAST STANDARD; PRT; 1790 AA.

AC P25386;

DT 01-MAY-1992 (Rel. 22, Created)

DT 01-MAY-1992 (Rel. 22, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Intracellular protein transport protein USO1.

GN USO1 OR INT1 OR YDL058W.

OS Saccharomyces cerevisiae (Baker's yeast).

OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.

OX NCBI_TaxID=4932;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=X2180-1A;

RX MEDLINE=91185402; PubMed=2010462;

RA Nakajima M., Hirata A., Ogawa Y., Yonehara T., Yoda K.,

RA Yanasaka M.;

RT "A cytoskeleton-related gene, usol, is required for intracellular

RT protein transport in Saccharomyces cerevisiae.";

RL J. Cell Biol. 113:245-260(1991).

RN [2]

RP SEQUENCE OF 782-1790 FROM N.A.

RA Hostetter M.K., Herman D.J., Bendel C.M., McClellan M., Tao N.,

RA Kendrick K.E.;

RL Submitted (FEB-1993) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 1-8 FROM N.A.

RA Bai Y., Symington L.S.;

RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: REQUIRED FOR PROTEIN TRANSPORT FROM THE ER TO THE GOLGI

CC COMPLEX.

CC -1- SUBCELLULAR LOCATION: CYTOPLASMIC. ASSOCIATED WITH INTRACELLULAR

CC MEMBRANES. PROBABLY PRESENT ON VESICLES OPERATIONAL BETWEEN THE

CC ER AND THE GOLGI COMPLEX.

CC -1- DOMAIN: THE RODLIKE TAIL SEQUENCE IS HIGHLY REPETITIVE, COMPOSED

CC OF AN HEPTAPEPTIDE REPEAT PATTERN CHARACTERISTIC OF ALPHA-HELICAL

CC COILED COILS. MAY FORM FILAMENTOUS STRUCTURES IN THE CELL.

CC -1- SIMILARITY: BELONGS TO THE VDP/USO1/YBL047C FAMILY.

CC -----

CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.igb-sib.ch/announce/>

CC or send an email to license@sib-sib.ch).

CC -----

DR EMBL; X54378; CAA38253.1; -

DR EMBL; L03188; AAB00143.1; -

DR EMBL; U53668; AAB66659.1; -

DR SGD; S0002216; USO1.

DR InterPro; IPR002017; Spectrin.

DR InterPro; IPR006955; Usol_p115_C.

DR InterPro; IPR006953; Usol_p115_head.

DR Pfam; PF04871; Usol_p115_C_1.

DR Pfam; PF04869; Usol_p115_head; 1.

KN Transport; Protein transport; Golgi stack; Cytoskeleton; Coiled coil.

FT DOMAIN 1 724

FT DOMAIN 725 1790

FT DOMAIN 465 487

FT DOMAIN 991 1790

FT DOMAIN 1172 1786

FT CONFLICT 847 847

FT CONFLICT 924 924

FT CONFLICT 1253 1253

FT CONFLICT 1319 1319

FT CONFLICT 1461 1461

FT CONFLICT 1581 1581

FT CONFLICT 1600 1600

FT CONFLICT 1661 1661

FT CONFLICT 1772 1772

D -> DEEDDBE (IN REF. 2).

SQ SEQUENCE 1790 AA; 206424 MW; 6CE2B216E9FD4818 CRC64;

Query Match 7.6%; Score 183.5; DB 1; Length 1790;

Best Local Similarity 20.3%; Pred. No. 0.15;

Matches 120; Conservative 91; Mismatches 200; Indels 179; Gaps 23;

QY 29 KSLHN--RELGAERPLNFIQAEAEEDKIKTYPPENKPGQSNYSFVDNLLKAI---T 82

DB 965 KSLANNYKMQAE--NESLIKAVEE-----SKNESSIQLSNLKNKIDMSQ 1008

QY 83 EKEKIEKERQIRSPDLNKLNVEDVSTKRNKLI-----DDYSTKSGLDHKKFOD--- 133

DB 1009 EKENFOIERGSIKIEKIQKKTISDLEQTBEBIISKSDSSKDEYESQISLLKEKLETTAT 1068

QY 134 -DPDGLHQLDGTPLTAEDIVHKIARIYEENDRAVDFDKIVSKLLNLGLITESQAH----- 187

DB 1069 ANDENVNKISBLTKTREETLEAEALAA--YKNLKNLELTLETSEKALKVEKNEEHKKEK 1126

QY 188 -TLEDEVAEVLQKLISKAN-----NYEED--PNKPTSWTNOAGKIPEK 229

DB 1127 IQLEKEATETKQQLNSLRANLESLEKEHEDLAAQLKKYEEQIANKEQYNE-EISQLNDE 1185

QY 230 VTPMAAIQDGLAKGENDETVSNLTLTNGLRRT-----KTYSEDN-- 270

DB 1186 ITSTQQENESIKK-KNDELEGEVRAKSTSEQSNLKKSEIDALNLQIKELKKKQNETNEA 1244

QY 271 -----FEELOYFPNP-----YALLKSIDSEKEAK 294

DB 1245 SLLESIKSVSETVKIKELQDECNPKFKEVSELEDKLKASEDKNSKYLELQK-ESEKIKE 1303

QY 295 EKETLIITIMKTLDIFVKNMVK-----YALLKSIDSEKEAK 1315

DB 1304 ELDAKTTTELKIQLEKTIINLSKAKSESELSRLKKTSSSEERKNAEEQLEKLNQIKNQ 1363

QY 316 -----YGTISP--EGVSYLENDEMIALQTKNLEKNATNISKLPAPSEK 361

DB 1364 AFEKERLLNNGSSITTOEYSEKINTLE--DELIRLQNELNKAKEIDNTRSELEKVSLS 1421

QY 362 SHEETDSTKEAAKMEKEVGSIKDS-TKDDNSNPGKTDPE-----KGTAYLEAIRK 414

DB 1422 NDELEEKQNTIKQLQDEILSYKOKITRNDKLSIERDNKRDLSEKLEQLRAAESKAK 1481

QY 415 NIEWLKHKDKGNKEDVDLSKVRDFINKQADAYVEKGILD-KEEAIAIKR 463

DB 1482 VEEGLKXLEESSKEAELEKSEKEMW-KKLESTIESNETELKSSMETIRK 1530

RESULT 7

MAPB RAT STANDARD; PRT; 2459 AA.

AC P15205; O62958; Q9ER21; Q9QW92;

DT 01-APR-1990 (Rel. 14, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 28-FEB-2003 (Rel. 41, Last annotation update)

DE Microtubule-associated protein 1B (MAP 1B) (Neuraxin) [Contains: MAP1

DE light chain LC1].

GN MAP1B.

OS Rattus norvegicus (Rat).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.

OX NCBI_TaxID=10116;

RN [1]

RP SEQUENCE OF 1-142 FROM N.A.

RC STRAIN=Sprague-Dawley; TISSUE=Testis;

RX MEDLINE=96257242; PubMed=8666295;

RA Liu D., Fischer I.;

RT "Isolation and sequencing of the 5' end of the rat microtubule-

RL associated protein (MAP1B)-encoding cDNA.";

RN Gene 172:307-308(1996).

RN [2]

RP SEQUENCE OF 96-2459 FROM N.A.; DOMAIN, AND INDUCTION.

RX STRAIN=Sprague-Dawley; TISSUE=Brain, and Glial tumor;

RX MEDLINE=92347374; PubMed=1639092;

RA Zauner W., Kratz J., Staunton J., Feick P., Wiche G.;
 RT "Identification of two distinct microtubule binding domains on
 RL recombinant rat MAP 1B.";
 RN Eur. J. Cell Biol. 57:66-74 (1992).
 [3]
 RP SEQUENCE OF 1541-2459 FROM N.A., AND TISSUE SPECIFICITY.
 RC TISSUE-Spinal cord;
 RX MEDLINE=90059871; PubMed=2555150;
 RA Rientz A., Greeningloh G., Hermans-Borgmeyer I., Kirsch J.,
 RA Littauer U.Z., Prior P., Gundelfinger E.D., Schmitt B., Betz H.;
 RT "Neuraxin, a novel putative structural protein of the rat central
 RT nervous system that is immunologically related to microtubule-
 RT associated protein 5.";
 RL EMOB J. 8:2873-2888 (1989).
 RN [4]
 RP DEVELOPMENTAL STAGE, AND PHOSPHORYLATION.
 RX MEDLINE=97405699; PubMed=9260743;
 RA Ma D., Nothias F., Boyne L.J., Fischer I.;
 RT "Differential regulation of microtubule-associated protein 1B (MAP1B)
 RT in rat CNS and PNS during development.";
 RL J. Neurosci. Res. 49:319-332 (1997).
 CC -!- FUNCTION: The function of brain MAPS is essentially unknown.
 CC Phosphorylated MAP1B may play a role in the cytoskeletal changes
 CC that accompany neurite extension. Possibly MAP1B binds to at least
 CC two tubulin subunits in the polymer, and this bridging of subunits
 CC might be involved in nucleating microtubule polymerization and in
 CC stabilizing microtubules.
 CC -!- SUBUNIT: 3 different light chains, LC1, LC2 and LC3, can associate
 CC with MAP1A and MAP1B proteins.
 CC -!- TISSUE SPECIFICITY: Nervous system (spinal cord, brain stem,
 CC cerebellum and cerebrum). Not expressed in liver, spleen, kidney,
 CC heart or muscle.
 CC -!- DEVELOPMENTAL STAGE: In cerebral cortex, spinal cord and sciatic
 CC nerve levels are high early in development but decrease during
 CC postnatal development and are low in adults. In dorsal root
 CC ganglia levels remain high throughout development.
 CC -!- INDUCTION: By nerve growth factor.
 CC -!- DOMAIN: Has a highly basic region with many copies of the sequence
 CC KKEE and KKEI/V, repeated but not at fixed intervals, which is
 CC responsible for the binding of MAP1B to microtubules.
 CC -!- PTM: LC1 is coexpressed with MAP1B. It is a polypeptide generated
 CC from MAP1B by proteolytic processing. It is free to associate with
 CC both MAP1A and MAP1B. It interacts with the amino-terminal region
 CC of MAP1B (by similarity).
 CC -!- PTM: Phosphorylated.
 CC -!- SIMILARITY: TO MAP1A.
 CC -!- CAUTION: A C-terminal fragment of this protein (residues 1597 to
 CC 2459) was originally described as neuraxin in Ref.3.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U52950; AAB17068.1; -;
 CC EMBL; X60370; CAC16162.1; -;
 CC EMBL; X16623; CAA34620.1; ALT_SEQ.
 CC PIR; A56577; A56577.
 CC InterPro; IPR000102; MAP1B neuraxin.
 CC Pfam; PF00414; MAP1B neuraxin; 10.
 CC PROSITE; PS00230; MAP1B_NEURAXIN; 8.
 KW Microtubules; Repeat; Phosphorylation.
 FT CHAIN ? 2459 MAP1 LIGHT CHAIN LC1.
 FT REPEAT 1869 1885 MAP1B 1.
 FT REPEAT 1886 1902 MAP1B 2.
 FT REPEAT 1903 1919 MAP1B 3.
 FT REPEAT 1920 1936 MAP1B 4.
 FT REPEAT 1937 1953 MAP1B 5.
 FT REPEAT 1954 1970 MAP1B 6.
 FT REPEAT 1988 2004 MAP1B 7.

REPEAT	2005	2021	MAP1B 8.
FT REPEAT	2022	2038	MAP1B 9.
FT REPEAT	2039	2055	MAP1B 10.
FT DOMAIN	559	1035	GLU-RICH
FT DOMAIN	588	786	LYS-RICH (HIGHLY BASIC, CONTAINS MANY KKEE AND KKEI/V REPEATS).
FT DOMAIN	2224	2312	LYS-RICH
FT CONFLICT	127	127	M -> V (IN REF. 1).
FT CONFLICT	140	140	T -> S (IN REF. 1).
FT CONFLICT	2112	2112	R -> K (IN REF. 3).
FT CONFLICT	2169	2169	L -> I (IN REF. 3).
SQ SEQUENCE	2459 AA; 269497 MW; 2E3F6872DEDB8BA2 CRC64;		

Query Match 7.5%; Score 180.5; DB 1; Length 2459;
 Best Local Similarity 22.7%; Pred. No. 0.31;
 Matches 120; Conservative 77; Mismatches 200; Indels 131; Gaps 22;

QY	22	KPGSQDQSLHNRSLAERPLNEQIAEEDKIK-KTYPPEKPGQSNYSFVNDLNLKKA	80
Db	549	KPLSSKSVRKSEKEEAPKASQVETPKVESKEKIVVKDKPKG-----VESKPS	600
QY	81	ITEKEIEKERSIRSSPLDNKLVNEDVDSTPNRKLDDVDSTKSLDHPQDDP--DGL	138
Db	601	VTEKEVSEKEEQSPVKAEEAKATESKPKVKDKVVKKEIKTKP--EKEKEPKKEVA	658
QY	139	HQDGTPLTADIVHKIAARIYEENDRAVFDKIVSKLNLGLITESQAHT-LEDEVAEVL	197
Db	659	KKEDKTPLKDKDEPKK-----EAKKEIKKEIKKEEKELKKEVKETPLDKAKKEV-	710
QY	198	QKLISKANNEYEDPNK-----PTSTWENQAG-----KIPEKVTPTMAA-	235
Db	711	KDDEKKEVKEEKEPKKEIKKISKDKIKKSTPLSDTKKPAALPKVAKKEEPTKKEPTAAG	770
QY	236	-IQD-GLAK-----GENDETVSNTL-----TLTNGLERRTK	264
Db	771	KLKDKGVKVIKKEGKTEAAATAGTAANAAGVAAAGAPAKLEAERSLMSSEPDITK	830
QY	265	TYSEDNFEELQYFPNFYALLKSIDSEKAE-----KETLITIMKTLIDFVKMMV	314
Db	831	DFEELKAEIDVAKDIKPOLLEIEDEKLEKTEPEGAUVIQTETEVS-----877	
QY	315	KYGTISPERGVSYLENDEMIALQTKNL---EKVATDNISKL-----PPAPSEKSHEE	365
Db	878	KGSAESDEGITTTEGEGE--CEQTPEELPFVEKQGVDDIEKFEDEGAGFESEAGDYE	935
QY	366	TDSTKEEAAKWEKYGSLKDKSTDDNSNPGKTDPEKGTAE--YLEAIRKNI-----416	
Db	936	EKAETEAEPEEDG-----EDNVSGSASKHSPTDEETAKAEADVHIKREKVSAGDDRA	992
QY	417	-----EWLKHKKKGNKEDYDLSKMRDFINKQADAVVEKGLDKKEAE	459
Db	993	EEDMDALEKGEARQSEEGEEES-----DKAEDAREDEHPDKTEAE	1035

RESULT 8
 RA50 METJA STANDARD; PRT; 1005 AA.
 ID RA50 METJA Q58718;
 AC Q58718;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR MJ1322.
 OS Methanococcus jannaschii.
 OC Archaea; Euryarchaeota; Methanococci; Methanococcales;
 OC Methanocaldococaceae; Methanocaldococcus.
 OX NCBI_TaxID=2190;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=JAL-1 / DSM 2661 / ATCC 43067;
 RX MEDLINE=96337999; PubMed=8688087;
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
 RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,

RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,
 RA Overbeek R., Kirchner E.F., Weinstock K.G., Merrick J.M., Glodek A.,
 RA Scott J.L., Geoghagen N.S.M., Weidman J.P., Fuhrmann J.L., Nguyen D.,
 RA Utterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
 RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.,
 FT "Complete genome sequence of the methanogenic archaeon, Methanococcus
 RT jannaschii.";
 RL Science 273:1058-1073(1996).
 CC -!- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mre11 complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mre11 by unwinding
 CC and/or repositioning DNA ends into the mre11 active site (By
 CC similarity).
 CC -!- SUBUNIT: Forms a complex with mre11 (By similarity).
 CC -!- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; U67572; AAB99331.1; -;
 DR TIGR; MJ1322; -;
 DR HAMAP; MF_00449; -;
 DR InterPro; IPR003593; AAA_ATPase.
 DR InterPro; IPR003439; ABC_transporter.
 DR InterPro; IPR003405; SMC_C.
 DR InterPro; IPR003395; SMC_N.
 DR Pfam; PF04423; Rad50_zn_hook; 1.
 DR Pfam; PF02483; SMC_C; 1.
 DR Pfam; PF02463; SMC_N; 1.
 DR ProDom; PD000006; ABC_transporter; 1.
 DR SMART; SM00382; AAA; 1.
 DR DNA_repair; Hydrolase; ATP-binding; Coiled coil; Complete proteome.
 KW NP_BIND 32 39
 FT DOMAIN 158 849
 FT COILED COIL (POTENTIAL).
 SQ SEQUENCE 1005 AA; 119387 MW; 9B8BB48173E788F3 CRC64;
 Query Match 7.5%; Score 179; DB 1; Length 1005;
 Best Local Similarity 22.3%; Pred. No. 0.12;
 Matches 116; Conservative 80; Mismatches 164; Indels 160; Gaps 27;
 QY 52 DKIKKTYPPENKQSNYSFVDNINLLKA-ITEKEIEKEKQSTRSSPLDNKLN-VEDVD 109
 Db 159 DEFEKCY--QKMGIEVKEVEKRLERIEGELNYKENYEKE-----LNKMSQLEE-- 205
 QY 110 STNKRLLDDVSTKSLGDLHKFDQDPLGHLQDGLTAEADIVHKAIRIYEENDRAVFD 169
 Db 206 --KNKKLMEINDKLNK-IKKEFEDIEKLFNEWENKLLYEFINKL-----EERKRALEL 257
 QY 170 KIVS-KLNLNGLITESQAH-----TLEDEVAEVLQKLSKEANNYEEDPNK 214
 Db 258 KNQELKILEYDLNVTVEARETLNRRHKDEYKYSKLVDRKIESRL--RELKSHVEDYLK 315
 QY 215 PTSTWENQAKIPKVTPTMAAIQDLAKGENDETVNTILTNGLERTKTYTSEDNFEEL 274
 Db 316 LTKQLEIKIGDI-EK-----LKEFINKSKYRDDIDNLDNLTKNI-----KDEIERV 360
 QY 275 QYFPNFVALLKSIDSEKEAKEK-----ETLITITMTKLIDFVKMKVKGTTISPE 322
 Db 361 ETIKLLEELNLENEEIEKIEKIKYRICECKEYVEKLEFEKAVENKUTLEYITL--- 417
 QY 323 EGVSYLENLDEMIALQPNKLEKNATD---NISKLPAPS-----EKSHEDTSTK-- 370
 Db 418 -----LQEKKSIEKNINDLETRINKLLETKNIDIESIENSLKEIEKKV 463
 QY 371 -----EBAAKMEKEYGSLKSTK-----DDNSNPGKK-----TDEPK-----GKT 405

Db 454 LENLQKEIENLKKLGEINSEIKRLKILDELKEVGKCPCKTPTIDENKKMELINQHK 523
 QY 406 E--AYLE-----AIRKNIEWLKKH-----DKKGNKEDYDL----- 433
 Db 534 QLNKKYTELEEEINKIRIEIKDEIKLKEIKDEENLKLTKLYLEKQSIQIELELKLKNY 563
 QY 434 SKMRDFINKQADAYVEKG-----ILD--KEBAEAIKRIYS 466
 Db 584 KEQLDEINKKISNYVINGKPVDEILEDIKSQLNKFKNFYN 623
 RESULT 9
 ID REST_HUMAN STANDARD; PRT; 1427 AA.
 AC P30622;
 DT 01-APR-1993 (Rel. 25, Created)
 DT 01-APR-1993 (Rel. 25, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Restin (Cytoplasmic linker protein-170 alpha-2) (CLIP-170) (Reed-
 DE Sternberg intermediate filament associated protein).
 GN RSN
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Peripheral blood monocytes;
 RX MEDLINE=92289675; PubMed=1600942;
 RA Bilbe G., Delabie J., Brueggen J., Richener H., Assebergs F.A.M.,
 RA Cerletti N., Sorg C., Odink K., Tarcsay L., Wiesendanger W.,
 RA de Wolf-Peters C., Shipman R.;
 RT "Restin: a novel intermediate filament-associated protein highly
 RT expressed in the Reed-Sternberg cells of Hodgkin's disease.";
 RL EMBO J. 11:2103-2113(1992).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92405160; PubMed=1356075;
 RA Pierre P., Scheel J., Rickard J.E., Kreis T.E.;
 RT "CLIP-170 links endocytic vesicles to microtubules.";
 RL Cell 70:887-900(1992).
 CC -!- FUNCTION: SEEMS TO BE A INTERMEDIATE FILAMENT ASSOCIATED PROTEIN
 CC THAT LINKS ENDOCYTIC VESICLES TO MICROTUBULES.
 CC -!- SUBCELLULAR LOCATION: CYTOPLASMIC, ASSOCIATED WITH THE
 CC CYTOSKELETON
 CC -!- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=2;
 CC Name=Long;
 CC IsoId=P30622-1; Sequence=Displayed;
 CC Name=Short;
 CC IsoId=P30622-2; Sequence=VSP_000765;
 CC -!- TISSUE SPECIFICITY: HIGHLY EXPRESSED IN THE REED-STERNBERG CELLS
 CC OF HODGKIN'S DISEASE.
 CC -!- SIMILARITY: Contains 2 CAP-Gly domains.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 CC EMBL; X64838; CAA46050.1; -;
 DR EMBL; M97501; AAA35693.1; -;
 DR PIR; S22695; S22695.
 DR Gene; HGNC:10461; RSN.
 DR MIM; 179838; -;
 DR GO; GO:0005768; C:endosome; TAS.
 DR GO; GO:0005882; C:intermediate filament; TAS.
 DR GO; GO:0015630; C:microtubule cytoskeleton; TAS.
 DR GO; GO:0008017; F:microtubule binding activity; TAS.

DR GO: GO:0006899; P-non-selective vesicle transport; TAS.
 DR InterPro; IPR000938; CAP_Gly.
 DR Pfam; PF01302; CAP_Gly; 2.
 DR SMART; SM00343; ZNF_C2HC; 1.
 DR PROSITE; PS00845; CAP_Gly 1; 2.
 DR PROSITE; PS0245; CAP_Gly 2; 2.
 KW Cytoskeleton; Microtubules; Coiled coil; Repeat; Alternative splicing.
 FT DOMAIN 78 120
 FT CAP_GLY 1.
 FT DOMAIN 143 204
 FT SER-RICH.
 FT DOMAIN 232 274
 FT CAP_GLY 2.
 FT DOMAIN 304 331
 FT SER-RICH.
 FT DOMAIN 350 1342
 FT COILED COIL (POTENTIAL).
 FT DOMAIN 1408 1421
 FT CCHC-BOX.
 FT VARSPLIC 457 491
 FT Missing (in isoform Short).
 FT /FTIC=VSP 000765.
 FT D -> E (IN REF. 2).
 FT CONFLICT 1069 1069
 FT SEQUENCE 1427 AA; 160989 MW; 0A4F166DD94254E8 CRC64;
 SQ
 Query Match 7.4%; Score 178.5; DB 1; Length 1427;
 Best Local Similarity 22.4%; Pred. No. 0.2;
 Matches 111; Conservative 79; Mismatches 182; Indels 123; Gaps 20;
 QY 34 RELSAE-----RPLNEQIAEAEEDKIKTYPENKPGQSNYSFVDNL-----NLLKA 80
 Db 769 RKASSEGSEMKLRQQL-EAAEKQIKHLEIKNAESSKASITRELQGRLEKLTNLQEN 827
 QY 81 ITE-----KEKIKERQ-----SIRSSPLD--NKL-----NVEDVDSTKN 113
 Db 828 LSEVSQVETLEKELQILKEKFAEAESEAVSQRSQETVKNLKHQKEQFNMLSDLEKL 887
 QY 114 RKLIDYDSTKGLDHKQDDPDGLHOLDGTPLTAEDIVHKIARIYENDRAVDFKIVS 173
 Db 888 RENLAD-----MEAKFREKDEREQL-----IKAKEKLEND-----IA 920
 QY 174 KLLNLGLITESQAHTLEDEV-----AEVLQKLISKEANNYE-----EDPNKPTISWTEN 221
 Db 921 EIMKSGNSGNSOUTKQNDLRLKRDVBELQULTKANENASFLQKSTEDMTVKAESQ 980
 QY 222 QAGKIPKVPMAAQDGLAGKENDETVNTLTNLGERTKTYSDENFELVFPNPFY 281
 Db 981 EAAKHGE-----EKKELERKLSL-----LEKWET-SHQCCQLK----- 1015
 QY 282 ALLKSIDSEKAKEKETITIMTKLIDFVKMMVKYGTISPEEGVSYLENLDEMIALQPN 341
 Db 1016 ARYERATSETKTHEIILQNLQKTLTDLTDKLG-----ARENSGLLQLEELRQADKA 1071
 QY 342 KLEKATNISKLPAPSEKSH--BETDSTKEAAKMEKEVGSGLKSDTKDNDNSPGGKTD 399
 Db 1072 KAAQTAEDAMQIMEQMTKEKTETLASLEDTKQTNKQLQNLDTLKNNL-KNVEELNKS 1130
 QY 400 EPKKGTEAYLEAIRKNIEWLKH-----DKGKNKEDYDLKMKRDFINKQADAYV 448
 Db 1131 ELLTVENQMEFREITLQKAAQKSQLQALQENVKLABELGRSDEVTSHQKLEE 1190
 QY 449 EKGILDKEEAEAIKR 463
 Db 1191 ERSVLNNLLEMKR 1205

RESULT 10
 ID RA50 PYRFU STANDARD; PRT; 882 AA.
 AC P58301.
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 15-SEP-2003 (Rel. 42, Last annotation update)
 DE DNA double-strand break repair rad50 ATPase.
 GN RAD50 OR PF1167.
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;

RN SEQUENCE FROM N.A., AND CHARACTERIZATION.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=2048553; PubMed=11029422;
 RA Hopfner K.-P., Karcher A., Shin D., Fairley C., Tainer J.A.,
 RA Carney J.P.;
 RT "Mrell and Rad50 from Pyrococcus furiosus: cloning and biochemical
 RT characterization reveal an evolutionarily conserved multiprotein
 RT machine.";
 RL J. Bacteriol. 182:6036-6041(2000).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome.";
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN X-RAY CRYSTALLOGRAPHY (1.6 ANGSTROMS) OF 1-149.
 RC STRAIN=Vcl / DSM 3638 / ATCC 43587 / JCM 8422;
 RX MEDLINE=20348838; PubMed=10892749;
 RA Hopfner K.-P., Karcher A., Shin D.S., Craig L., Arthur L.M.,
 RA Carney J.P., Tainer J.A.;
 RT "Structural biology of Rad50 ATPase: ATP-driven conformational
 RT control in DNA double-strand break repair and the ABC-ATPase
 RT superfamily.";
 RL Cell 101:789-800(2000).
 RL CC
 CC -1- FUNCTION: Involved in DNA double-strand break repair (DSBR). The
 CC rad50/mrell complex possesses single-strand endonuclease activity
 CC and ATP-dependent double-strand-specific exonuclease activity.
 CC Rad50 provides an ATP-dependent control of mrell by unwinding
 CC and/or repositioning DNA ends into the mrell active site.
 CC -1- SUBUNIT: Forms a complex with mrell.
 CC -1- SIMILARITY: BELONGS TO THE SMC FAMILY. RAD50 SUBFAMILY.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; A010225; AAL81291.1; -
 DR PDB; 1F2T; 20-SEP-00.
 DR PDB; 1F2U; 02-AUG-00.
 DR PDB; 1I18; 30-MAY-01.
 DR PDB; 1L8D; 28-AUG-02.
 DR HAMAP; MF_00449; -; 1.
 DR InterPro; IPR003439; ABC transporter.
 DR InterPro; IPR003395; SMC N.
 DR Pfam; PF04423; Rad50 zn hook; 1.
 DR Pfam; PF02463; SMC N; 1 hook; 1.
 DR ProDom; PD000006; ABC transporter; 1.
 DR DNA repair; Hydrolase; ATP-binding; Coiled coil; 3D-structure;
 DR Complete proteome.
 DR NP BIND 30 37 ATP.
 DR NP BIND 148 744 COILED COIL (POTENTIAL).
 DR DOMAIN 882 AA; 103839 MW; 3ADCBD250382A99E CRC64;
 SQ SEQUENCE 882 AA; 103839 MW; 3ADCBD250382A99E CRC64;
 Query Match 7.4%; Score 178; DB 1; Length 882;
 Best Local Similarity 22.1%; Pred. No. 0.12;
 Matches 112; Conservative 90; Mismatches 172; Indels 132; Gaps 21;
 QY 39 ERLNEQIAEAE--DKIK---KTYPPENKPGQSNYSFVDNLNLLKA-----ITEKEKIE 88
 Db 260 KRTLEERIKNTTEYLEKLEKEKELESEQVEITSIKKVDVAYLALKEFKNEYLDKKYKIE 319
 QY 89 KERQSIRSSPLDNKLNVEDVDSTKNRKLIDYDSTKGLDHKQDDPDGLHOLDGTPLTA 148
 Db 320 KE--LITVEELINEIQ-----KRIEELNEKESEKEKLENE-----K 353
 QY 149 EDIVHKIARIYENDRAVDFKIVSKLLNLGLITESQAHTLEDEVAVELQKLISKEANNY 208
 Db 149 EDIVHKIARIYENDRAVDFKIVSKLLNLGLITESQAHTLEDEVAVELQKLISKEANNY 208


```

Db 354 KEILNKLAIL---EKDQHYBEIKAKKENLQKKEKLGKSPEDIKKLEEBLETKK-TTI 409
QY 209 EEDPNKPTSWTENQAKIPEKVTMAAIOQGL--AKGS-----NDE----- 247
Db 410 EERNEIT---QRIGELKNIGDLKTAIEELKAKGKCPVCGRELTDHREELLSKYHL 465
QY 248 -----TVSNLTLTNGLERTKTYSEDNFEELQYFNFYALLKSIDS-----E 290
Db 466 DLNNSKNTLAKLIDRKSELERLRIDME-IKRLTPLLTVAEQIRSEELNVNVLKIE 524
QY 291 KEAKEKETLIITIMKTLDIFVQWVKYGTISPEBGVSYLENDEMIA--LQTKNKLKQAT 348
Db 525 KNATEYKELBELRTLEGRIRGLA-----EDLKKLAPLEKKLAALIHKKQLEKELK 576
QY 349 DNISKLPAPSEKSHETDSTKEEAAKMEKEYGSLKSTKD-----DNSNPGKTD 399
Db 577 ELNTKL-ESFGFKSVEDLDSKLELEIYKYLYTLNLSKKELEITQREIAKAKETLMSF 635
QY 400 EPKGTETAYLEAIRKNIEWLKKHDKGNKEDY-----DLSKMR 437
Db 636 BELAEVADIERIEKLSQLK---QKYNEEYKKKREKEKELEKELARLEAQKKELEKRR 692
QY 438 DFINKQADAVEKGILDKEEAIAKR 463
Db 693 DTIK-----STLEKKAENKERNRVKK 714

RESULT 11
BAG STRAG
ID BAG STRAG STANDARD; PRT; 1164 AA.
AC P27951;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE IGA FC receptor precursor (Beta antigen) (B antigen).
GN BAG.
OS Streptococcus agalactiae.
OC Bacteria; Firmicutes; Lactobacillales; Streptococcaceae;
OC Streptococcus.
OX NCBI_TaxID=1311;
RN 11
RP SEQUENCE FROM N.A., AND SEQUENCE OF 38-48.
RC STRAIN=LA239;
RX MEDLINE=91312121; PubMed=1857207;
RA Jerlstrom P.G., Chhatwal G.S., Timmis K.N.;
RA Bateman A., Eddy S.R., Chothia C.;
RT "The Iga-binding beta antigen of the c protein complex of Group B
RT streptococci: sequence determination of its gene and detection of two
RT binding regions."
RL Mol. Microbiol. 5:843-849(1991).
RN [2]
RP IDENTIFICATION OF IG-LIKE DOMAIN.
RX MEDLINE=97035265; PubMed=8880921;
RA Bateman A., Eddy S.R., Chothia C.;
RT "Members of the immunoglobulin superfamily in bacteria."
RL Protein Sci. 5:1939-1942(1996).
CC -1- SUBCELLULAR LOCATION: Attached to the cell wall peptidoglycan by
CC an amide bond (potential).
CC -1- SIMILARITY: Contains 1 immunoglobulin-like domain.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; X59771; CAA2442.1; -.
CC PIR; S15330; FCSOAG.
DR InterPro; IPR004829; Csurface antigen.
DR InterPro; IPR005877; Gpos YSIRK.
DR InterPro; IPR001899; Gram_pos_anchor.
DR InterPro; IPR003599; Ig.

```

```

DR InterPro; IPR006192; LPXTG.
DR Pfam; PF00746; Gram_pos_anchor; 1.
DR Pfam; PF05062; RICH; 1.
DR Pfam; PF04650; YSIRK_signal; 1.
DR Pfam; PF053432; Csurface_antigen; 1.
DR Pfam; PF053432; Csurface_antigen; 1.
DR SMART; SM00409; IG; 1.
DR TIGRFAMs; TIGR01167; LPXTG_anchor; 1.
DR TIGRFAMs; TIGR01168; YSIRK_signal; 1.
DR PROSITE; PS00847; GRAM_POS_ANCHORING; 1.
DR PROSITE; PS00835; IG_LIKE_FALSE_NEG.
DR Cell wall; Peptidoglycan-anchor; Receptor; Repeat; Signal;
KW Immunoglobulin domain.
FT SIGNAL 1 37
FT CHAIN 38 1135 IGA FC RECEPTOR.
FT PROPEP 1136 1164 REMOVED BY SORTASE (POTENTIAL).
FT DOMAIN 434 534 IG-LIKE.
FT DOMAIN 199 438 IGA-BINDING (POTENTIAL).
FT DOMAIN 439 826 IGA-BINDING (POTENTIAL).
FT DOMAIN 827 945 PRO-RICH REPEATS.
FT SITE 1132 1136 LPXTG SORTING SIGNAL (POTENTIAL).
FT MOD_RES 1135 1135 AMIDE-LINKED TO CELL WALL (POTENTIAL).
FT SEQUENCE 1164 AA; 131051 MW; 65DE94AF720A5474 CRC64;

Query Match 7.2%; Score 172.5; DB 1; Length 1164;
Best Local Similarity 21.6%; Pred. No. 0.3;
Matches 124; Conservative 77; Mismatches 189; Indels 185; Gaps 25;

QY 36 LSAERPLNE-----QIAEAEEDKIKKTYPENKPGOSYFVDNLLMLLKAIT-EKEK 86
DB 149 LELENQFNETNRLHLIKQHEEVEKOKKAK-----QOKTLKQSDTKVDLSNIDKELNHSQ 204
QY 87 IEK--EROSIRSSPLDNKL-----NVEDVDSTKNRKLIDYDSTKSGLDHK 130
DB 205 VEKMAEQKGTINEDKXSMKLIEDIRKQAOADKEDAEVKVEELGKLFSTKAGLDQ 264
QY 131 FQDDPDGLHQLDGTPLTADIVHKI-----AARIYEENDRAVDFKIVSKLLNLG 179
DB 265 IQE-----HVKKET--SSEENTQKVDHYANSQNLAKSLEELDKATTNEQATQVKQF 317
QY 180 L-----ITESAHTLEDEVAEV-----LQKLISKE---ANNYEE 210
DB 318 LENAQKLEIOPLIKETNVKLYKAMSLEQVEKELHENSEANLEDLVAKSKEIVREYEG 377
QY 211 DPNKPTSWTE-----NOAGKPIPEKVTMAAIOQGLAKGENDETVPN 251
DB 378 KLNQSKNLPKQLBEEAHKLVQVVEDFRKKFKTSEQVTPKRVKRDLANENNQ---Q 434
QY 252 TLTTLNGLERTKTYSEDNFEELQYFNFYALLKSIDSKEAKEKETLITIMKTLDIFVK 311
DB 435 KIELTVSPENITVYEGED-----VKFTVTAKS-DS-----KTLDFSD 471
QY 312 MMVYKG-TISPEGVSYLENLD-----EMIALQTK-----NKLEKNATD 349
DB 472 LUTKYNPSVDRISTNYKTNTDNHKAIBITIKNLKLNESQVTLKAKDDSGNVVEKFTFI 531
QY 350 NISKLPAPSEKSHETDSTKEEAA-----KMEKEYGSLKD 385
DB 532 TVQKKEEQVKTPEQKDSKTEEVQPEKPNQKQNLQELIKSAQOELEKELKELME 591
QY 386 STKDDNSNPGKTDPEPKKTEYALEAIRKNIEWLKK-----HDKGNKEDYDLSKMRDFTN 441
DB 592 Q-PEIPSNP--EYGIQKSIWESQKEPIQEAITSFKKIIGDSSSKYTYTHYFNKYKSDPMN 648
QY 442 KQADAYVEKGILDKEEAEB-----AIKEIYSS 467
DB 649 YOLHAQME--MLTRKVVQVMNKNYPDNABEIKKIPES 681

RESULT 12
TRDN CANFA
ID TRDN CANFA STANDARD; PRT; 700 AA.
AC P82179;
DT 28-FEB-2003 (Rel. 41, Created)

```


28-FEB-2003 (Rel. 41, Last sequence update)
 28-FEB-2003 (Rel. 41, Last annotation update)
 Triadin.
 TRDN.
 Canis familiaris (Dog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
 NCBI_TaxId=9615;
 OX [1]
 RN SEQUENCE FROM N.A., AND ALTERNATIVE SPLICING.
 RP TISSUE=Heart, and Skeletal muscle;
 RC MEDLINE=94428545; PubMed=10497235;
 RX Kobayashi Y.M., Jones L.R.;
 RA "Identification of triadin 1 as the predominant triadin isoform
 RT expressed in mammalian myocardium.";
 RL J. Biol. Chem. 274:28660-28668(1999).
 CC -I- FUNCTION: MAY BE INVOLVED IN ANCHORING CALDESQUESTRIN TO THE
 CC JUNCTIONAL SARCOPLASMIC RETICULUM AND ALLOWING ITS FUNCTIONAL
 CC COUPLING WITH THE RYANODINE RECEPTOR (BY SIMILARITY).
 CC -I- SUBCELLULAR LOCATION: Type II membrane protein. Sarcolemmal
 CC reticulum.
 CC -I- ALTERNATIVE PRODUCTS:
 CC Event=Alternative splicing; Named isoforms=3;
 CC Names=Skeletal;
 CC IsoId=P82179-1; Sequence=Displayed;
 CC Names=Cardiac 1;
 CC IsoId=P82179-2; Sequence=VSP_004001, VSP_004002;
 CC Names=Cardiac 3;
 CC IsoId=P82179-3; Sequence=VSP_004003, VSP_004004;
 CC Names=Cardiac 3;
 CC -I- TISSUE SPECIFICITY: SKELETAL AND CARDIAC MUSCLE.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation at
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial/
 CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
 CC or send an email to license@isb-sib.ch).

 DR EMBL; AF165916; AAF00222.1; -.
 DR EMBL; AF165915; AAF00221.1; -.
 DR EMBL; AF165917; AAF00223.1; -.
 KW Transmembrane; Sarcolemmal reticulum; Glycoprotein;
 KW Alternative splicing.
 FT INIT MET 0
 FT FT 1 46
 FT FT DOMAIN 47 67
 FT TRANS MEM 68 700
 FT FT DOMAIN 74 74
 FT CARBOHYD 616 616
 FT CARBOHYD 257 277
 FT VAR SPLIC
 FT FT 278 700
 FT VAR SPLIC
 FT FT 466 466
 FT VAR SPLIC
 FT FT 524 579
 FT VAR SPLIC
 FT FT 700 AA; 78152 MW; F033E3A1BE0C56 CRC64;
 FT FT 7.0%; Score 169; DB 1; Length 700;
 FT FT 20.0%; Pred. No. 0.24;
 FT FT 97; Conservative 75; Mismatches 176; Indels 136; Gaps 19;
 FT FT 27 QDKSLH----NRRLSAPRLNEQIAEEBDEKIKKTYPPENKPGSNYSFDNLLKAIT 82
 FT FT 139 KQEKIHKEAKEEKPERKTLAKVAHKVEKSEKK-----AT 181

Qy	83	EKEKIEKQIRSPDLNKLNVEDVDSTKRNKLIDDDYDSTKSLGDLHFKQDDPDGLHQLD	144
Db	182	HKEKIEKEK-----ETKTMAKEERAKTEEKIK--KEVKGGRQEKVKPTAAKVKEVQ	233
Qy	143	GTPLTAEDIVHKIARI--YEENDRAVFDKIVSKLLNLGLITESQAHTL-----	189
Db	234	KTPPKAKEKEKETAAVAHKKQDOYAFCRYIMDMFVHGDURPGQSPALPPPLPTVQASR	293
Qy	190	-----EDEVAEVLQKLISKEANNYBEDPNKPTS-----WTENQAKIPE---	228
Db	294	PTPASPTLEGKEEBEKKKAEKVTSETKKEKEDVKKSKDKDTAIDVEKKEFGKAPETKQ	353
Qy	229	---KVTMAAIQDGLAKGENDETVSNLTULTNGLERRTKTYSEDNFELOYPFNFYALLK	285
Db	354	GTIVVAQAAK-----KDEKEDSKTKT-----PVBEHPKGGKQEKKE-----K	394
Qy	286	SIDSEKAKEKETLITIMKTLIDFVQMVKYGTITSPESGVSYLENLDEMIALQTKNKLK	345
Db	395	YVEPAKSSKKEHS-----APSE-----KQVKAKTERAKEE	424
Qy	346	NATDNISKLFPAPSEKSHETDSTKEEAAKMEKEYGSLKDSKODNSNPGGKTDE--PKG	403
Db	425	TSAASTKAVPG--KKEETKTVEQEIREEK---SGKTSTASKDKEPEIKKDEKMPKA	478
Qy	404	KTEAYLEAIRKNIEWLKHD---KGNKEDYDLGKMRDFIN-----KQADAYVEKG	451
Db	479	DKEVKPPQSQVKEEKESQVKKEAKPEQDIAKPEKTVSHGKPEEKVVKQVKA-TEKA	537
Qy	452	ILDK 455	
Db	538	AIEK 541	

RESULT 13

KP5C_HUMAN

ID

KP5C_HUMAN

STANDARD;

PRT;

957 AA.

AC

O60282; O95079;

DT

16-OCT-2001 (Rel. 40, Last sequence update)

DT

16-OCT-2001 (Rel. 40, Last annotation update)

DE

Kinesin heavy chain isoform 5C (Kinesin heavy chain neuron-specific 2).

DE

KIF5C OR NKHC2 OR KIAA0531.

GN

Homo sapiens (Human)

OS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX

NCBI_taxid=9606;

[1]

SEQUENCE FROM N.A.

RP

TISSUE=Brain;

RC

MEDLINE=98290545; PubMed=9628581;

RA

Nagase T., Ishikawa K.-I., Miyajima N., Tanaka A., Kotani H.,

RA

Nomura N., Ohara O.;

RA

"Prediction of the coding sequences of unidentified human genes. IX. The complete sequences of 100 new cDNA clones from brain which can code for large proteins in vitro.";

RT

DNA Res. 5:31-39(1998).

RL

[2]

RN

SEQUENCE OF 355-585 FROM N.A.

RP

Engelender S., Sharp A.H., Colomer V., Tokito M.K., Lanahan A.,

RA

Worley P., Holzbaur E.L.F., Ross C.A.;

RT

"Huntingtin associated protein 1 (HAP1) interacts with the p150Glued subunit of dynactin.";

RL

Submitted (JUN-1997) to the ENBL/GenBank/DBJ databases.

CC

-1- FUNCTION: KINESIN IS A MICROTUBULE-ASSOCIATED FORCE-PRODUCING PROTEIN THAT MAY PLAY A ROLE IN ORGANELLE TRANSPORT.

CC

-1- SUBUNIT: OLIGOMER COMPOSED OF TWO HEAVY CHAINS AND TWO LIGHT CHAINS.

CC

-1- TISSUE SPECIFICITY: HIGHEST EXPRESSION IN BRAIN, PROSTATE, AND TESTIS, AND MODERATE EXPRESSION IN KIDNEY, SMALL INTESTINE, AND OVARY.

CC

-1- DOMAIN: COMPOSED OF THREE STRUCTURAL DOMAINS: A LARGE GLOBULAR N-TERMINAL DOMAIN WHICH IS RESPONSIBLE FOR THE MOTOR ACTIVITY OF

CC KINESIN (IT HYDROLYZES ATP AND BINDS MICROTUBULE), A CENTRAL
 CC ALPHA-HELICAL COILED COIL DOMAIN THAT MEDIATES THE HEAVY CHAIN
 CC DIMERIZATION; AND A SMALL GLOBULAR C-TERMINAL DOMAIN WHICH
 CC INTERACTS WITH OTHER PROTEINS (SUCH AS THE KINESIN LIGHT CHAINS),
 CC VESICLES AND MEMBRANOUS ORGANELLES.
 CC -1- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY. KINESIN
 CC SUBFAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AB011103; BAA25457.1; -;
 CC EMBL; AF010146; AAD01436.1; -;
 CC HSSP; P56536; 2KIN.
 CC Genew; HGNC:6325; KIF5C.
 CC MIM; 604593; -;
 CC GO; GO:0005871; C:kinesin complex; TAS.
 CC GO; GO:0003777; F:microtubule motor activity; TAS.
 CC GO; GO:0006996; P:organelle organization and biogenesis; TAS.
 CC InterPro; IPR001752; kinesin_motor.
 CC Pfam; PF00225; kinesin; 1.
 CC PRINTS; PR00380; KINESINHEAVY.
 CC SMART; SM00129; KISC; 1.
 CC PROSITE; PS00411; KINESIN MOTOR DOMAIN1; 1.
 CC PROSITE; PS00607; KINESIN MOTOR DOMAIN2; 1.
 CC Motor protein; Microtubules; ATP-binding; Coiled coil.
 CC DOMAIN 1 386 KINESIN-MOTOR (BY SIMILARITY).
 CC DOMAIN 406 923 COILED COIL.
 CC DOMAIN 859 956 GLOBULAR.
 CC DOMAIN 174 315 MICROTUBULE-BINDING.
 CC NP_BIND 86 93 ATP (BY SIMILARITY).
 CC CONFLICT 355 360 TLKVI -> STHAVI (IN REF. 2).
 CC CONFLICT 583 585 EFT -> DRV (IN REF. 2).
 CC SEQUENCE 957 AA; 109494 MW; A9F25BBIC994322A CRC64;

Query Match 7.0%; Score 167.5; DB 1; Length 957;
 Best Local Similarity 20.3%; Pred. No. 0.42;
 Matches 103; Conservative 98; Mismatches 182; Indels 125; Gaps 23;
 QY 25 GSQDKSLHNR-----ELSAERPLNEQIAEAEEDKIKKTPPENKPGQSNYSFVDNLNL- 78
 DB 321 GQAKTKNTVSVNLEITAE-----EWKKYKEKEKNTLNKVIQHLMEEL 367
 QY 79 -----KAITEKEKIEKERSIRSSPLDNKLVNEDVD-----STKNRKLIDDYDSTKSG 126
 DB 368 NRWRNGEAVPEDEQISAKDQK-NLEPCDNTPIIDNAPVAVGISTEEK---EKYDEEISS 423
 QY 127 LDHKFQDDPGLHQLDGTPLTAEDIVHKIARIYEEND-----RAVFDKIVSKLNLNGLI 181
 DB 424 LYQLDLDKDEINQ-----QSOLAELKQOQMLDDELLASTRDYKIEQIEELTRLQIE 476
 QY 182 TESQAHTLEDEAVELQKLISKEANNYE---EDPNKPTSWTENQAGKPEKVTPEMAIAD 238
 DB 477 NEA-----AKDEVKEVLQAL-BELAVNYDQKSQVEDKTRANEQUTDELAKQITLTITTOR 531
 QY 239 GLAK-----GENDETSTNLTITLNGLER-----TK 264
 DB 532 ELSQQLSBNHQKRAATEILNLLKDLGEIGGIITND---VKTLADVNGVIEEFTMAR 588
 QY 265 TYSEDNFEELQYFNFALLKS--IDSEKAKEKETLITMTKTLIDFKVMVKGTISPE 322
 DB 589 LYISKMKSEKSVLNRSEKLESQMSQDNRRKNASERELAAQCLLSQHEAKIKSLT--- 644
 QY 323 EGVSYLENLDEMIALQTKNLEK---NATDNISKLPFAPSEKSGHEEDVSTKE----- 371
 DB 645 ---DYMQNM-----QKRQLESQDLSBELAKL--RAQEKHVEFSQDKEKHLTRLQ 694
 QY 372 EAAWMEKEYGSLKDDSTKDDNSNPGGKTDEPKGKTEAYLEAIRKNIEWLKKHKDKGNKEDY 431

Db 695 DAEMKKALEQQMESHRAHQKLSRLDETEEKQKIIDEIRDLNQLQLEQEK-LSSDY 753
 QY 432 DLSKWRDPFINKQADAYVEKGIL--DKEE 457
 Db 754 NKLKIED--QEREMKLEKLLLLNDKRE 778

RESULT 14
 ID MLPI YEAST STANDARD; PRT; 1875 AA.
 AC Q02455;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-JUN-1994 (Rel. 29, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Myosin-like protein MLPI.
 GN MLPI OR YKR095W OR YKR415.
 OS Saccharomyces cerevisiae (Baker's yeast).
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
 OX NCBI_TaxID=4932;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=S288C;
 RX MEDLINE=93247549; PubMed=8483450;
 RA Kosling R., Nguyen T., Chen E.Y., Botstein D.;
 RT "A new yeast gene with a myosin-like heptad repeat structure.";
 RL Mol. Gen. Genet. 237:359-369(1993).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=94205265; PubMed=8154186;
 RA Bou G., Esteban P.F., Baladron V., Gonzalez G.A., Cantalejo J.G.,
 RA Remacha M., Jimenez A., del Rey F., Ballesta J.P.G., Revuelta J.L.;
 RT "The complete sequence of a 15,820 bp segment of Saccharomyces
 RT cerevisiae chromosome XI contains the UBI2 and MPL1 genes and three
 RT new open reading frames.";
 RL Yeast 9:1349-1354(1993).
 CC -1- FUNCTION: MYOSIN-LIKE PROTEIN THAT IS PROBABLY INVOLVED IN DNA
 CC REPAIR.
 CC -1- SIMILARITY: SOME, TO THE TPR ONCOGENE.
 CC -1- CAUTION: REF.2 MISQUOTES THE GENE NAME AS "MPL1".
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L01992; AAA34783.1; -;
 CC EMBL; X73541; CAAS1948.1; -;
 CC EMBL; Z28320; CAA82174.1; -;
 CC PIR; S38173; S38173.
 CC SGD; S0001803; MLPI.
 CC GO; GO:0005635; C:nuclear membrane; IDA.
 CC GO; GO:0005854; C:nucleoplasm; IDA.
 CC GO; GO:0006606; P:protein-nucleus import; IDA.
 CC Coiled coil; DNA repair.
 KW DOMAIN 69 487 COILED COIL (POTENTIAL).
 FT DOMAIN 531 1678 COILED COIL (POTENTIAL).
 FT DOMAIN 1834 1866 COILED COIL (POTENTIAL).
 FT CONFLICT 301 301 R->A (IN REF. 1).
 SQ SEQUENCE 1875 AA; 218455 MW; 683A0D34C9066867 CRC64;

Query Match 6.9%; Score 166.5; DB 1; Length 1875;
 Best Local Similarity 23.6%; Pred. No. 1;
 Matches 112; Conservative 85; Mismatches 178; Indels 99; Gaps 28;
 QY 30 SLNRELASRPNEQIAEAEEDKIKKTPPENKPGQSNYSFVDNLNLKAIKE-KEKIE 88
 DB 588 TLKSEKMDLESRIEELQLELEE--LKTSTVPNE-----ASYSNV-----TIKQLTETKRDLE 637

QY 89 KERQSIKSPNDKLVNEDVDSTKRNKLI-----DDYDSTKSGLDHKKQDDPDGLHQLD 143
 Db 638 SQVQD-----LQTRISQITRESTENSLNKEIQDLYDS-KSDISIK-----LGKEKS 684
 QY 144 TPLTAED---IVHKIAARIYEEND--RAVFOKIVSKLNLGLITVESQAHTLEDEVAEVLQ 198
 Db 695 SRILAEERFKLLSNTLDLTKAENDQLRKRFYDLQNTILK-----QDSKTHETILNEVSVCKS 740
 QY 199 KL--ISKANNYEEDPNKPTSWTEN---OAGKI-PEK-----VTPMAAIQ-----DGLAK 242
 Db 741 KLSIVETELLNKEKQKRLVHLEKNLQKLNKLSPEKSLRIMVTQLQTLQKEREDLLEE 800
 QY 243 -----GENDETVSNLTILTNGLERTKYTSEDNFEELQYFPNFYALLKSDISEKEA 293
 Db 801 TRKSCQKIDLEALSSELKETSQKHIIHQLEEDNNSNTWQNKTEALK-----852
 QY 294 KEKETLIT-INKTLTDFVKMVMKYGTISPVEGVSYLENLDEMIALQTKNLEKKNATONIS 352
 Db 853 KYESVITSVDSKQTDIEKLQYKVSLEKE-----IE--EDKIRLHTYVWDEITINDSL 905
 QY 353 KLFAPAPSEKSHETDSTKKEAAKMEKYGSLKSDTKDSDNSNPGKTDPEKGTAYLEAI 412
 Db 906 R-----KELEKSKINLTDAYSQIKYKDLVETTSQSLQOTNSKLDLDE--SFKDFTNQI 955
 QY 413 RKNIEWLKHDKGNKED-YDLSKMRDP-INKQADAYVEKGLDKKEEAIAIKRI 464
 Db 956 -KNIT-----DEKTSLEDKISLLSQMFNANNELD--LQKGMKEKADFKRI 1001

RESULT 15

CENE_HUMAN
 AC Q02224; STANDARD; PRT; 2663 AA.
 DT 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 30-MAY-2000 (Rel. 39, Last annotation update)
 DE Centromeric protein E (CENP-E protein).
 GN CENPE.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP MEDLINE=93024922; PubMed=1406971;
 RA Yen T.J., Li G., Schaar B.T., Szilak I., Cleveland D.W.;
 RT "CENP-E is a putative kinetochore motor that accumulates just before
 mitosis";
 RL Nature 359:536-539(1992).
 RN [2]
 RP CHARACTERIZATION.
 RA Thrower D.A., Jordan M.A., Schaar B.T., Yen T.J., Wilson L.;
 RT "Mitotic HeLa cells contain a CENP-E-associated minus end-directed
 microtubule motor";
 RL EMBO J. 14:918-926(1995).
 RN [3]
 RP CHARACTERIZATION.
 RX MEDLINE=98437347; PubMed=9763420;
 RA Chan G.K.T., Schaar B.T., Yen T.J.;
 RT "Characterization of the kinetochore binding domain of CENP-E reveals
 interactions with the kinetochore proteins CENP-F and HUBRI";
 RL J. Cell Biol. 143:49-63(1998).
 CC -!- FUNCTION: MINUS-END DIRECTED MICROTUBULE MOTOR. PROBABLE
 KINETOCORE MOTOR. ACCUMULATES JUST BEFORE MITOSIS AT THE G2 PHASE
 OF THE CELL CYCLE. PROBABLY IMPORTANT FOR CHROMOSOME MOVEMENT
 AND/OR SPINDLE ELONGATION.
 CC -!- SUBUNIT: INTERACTS WITH CENP-F AND HUBRI KINASE.
 CC -!- SUBCELLULAR LOCATION: ASSOCIATES WITH KINETOCORES DURING
 CONGRESSION, RELOCATES TO THE SPINDLE MIDZONE AT ANAPHASE, AND IS
 QUANTITATIVELY DISCARDED AT THE END OF THE CELL DIVISION.
 CC -!- SIMILARITY: BELONGS TO THE KINESIN-LIKE PROTEIN FAMILY.
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; Z15005; CAA78727.1; --
 DR PIR; S28261; S28261.
 DR HGSP; P17119; 3KAR.
 DR Genew; HGNC:1856; CENPE.
 DR GK; Q02224; --
 DR MIN; 117143; --
 DR GO; GO:0005699; C:kinetochore; TAS.
 DR GO; GO:0005634; C:nucleus; TAS.
 DR GO; GO:0008350; F:kinetochore motor activity; TAS.
 DR GO; GO:000067; P:DNA replication and chromosome cycle; TAS.
 DR GO; GO:0007079; P:mitotic chromosome movement; TAS.
 DR GO; GO:0007080; P:mitotic metaphase plate congression; TAS.
 DR InterPro; IPR001752; kinesin_motor.
 DR Pfam; PF00225; kinesin; 1.
 DR PRINTS; PR00380; KINESINHEAVY.
 DR SMART; SM00129; KISC; 1.
 DR PROSITE; PS00411; KINESIN MOTOR DOMAIN; 1.
 DR PROSITE; PS0067; KINESIN MOTOR DOMAIN2; 1.
 KW Motor protein; Cell division; ATP-binding; Coiled coil; Mitosis;
 KW Cell cycle; Centromere.
 FT DOMAIN 1 335 KINESIN-MOTOR
 FT DOMAIN 336 2471 COILED COIL (POTENTIAL).
 FT DOMAIN 2472 2663 GLOBULAR (POTENTIAL).
 FT NP BIND 86 93 ATP (BY SIMILARITY).
 SQ SEQUENCE 2663 AA; 312087 MW; CEFCL3880C8C8C8 CRC64;
 Query Match 6.9%; Score 166; DB 1; Length 2663;
 Best Local Similarity 21.8%; Pred. No. 1.7;
 Matches 119; Conservative 103; Mismatches 175; Indels 150; Gaps 28;
 QY 26 SQKSLNRLSRAERPLN---EQIAEAEEDKIKKTPPENKPGQSNYSFVDNLLNLLKAIT 82
 Db 1102 AQEK---NHAIKKEGELSRFCDLAEVEEKLKESQQLQKQQL-----LNQVEEMS 1151
 QY 83 EKKEIKERQSIKSPNDKLVNEDVDSTK---NRKLIDDYDSTKS-----GLDHK 130
 Db 1152 EMQKKINEIENLKNELKNELTLEHMETELELAQKLNENVEVKSTIKERKVLKELQKS 1211
 QY 131 FQDDPDGL-----HQLDGTPL-TAEDI-----VH-KIAARIYEENDRAVDFKIVSKLNLGL 180
 Db 1212 FETERDHLRGYIREIEATGLQTKELKIAHILKHEQETIDELRRSVSEK-TAQIINTQD 1270
 QY 181 ITESQAHTLEDEVAEVL---QKLIS--KEANNYEEDDNKPTSWTENQAGKIPKVTPTMAA 235
 Db 1271 LEKS--HTKLOEIPVLHBEQELLPNVKVSETQETWNELELLTEQSTTK---DSTTLAR 1325
 QY 236 IQ-----DGLAK-GENDETIVSNTL-----TLTNGLERRTK 264
 Db 1326 IEMERLRLNEKFOESQBEIKSLTKERDNLKTIKEALEVHDKLKEHRETLAKIQESQSK 1385
 QY 265 TYSEDNPEEL-----QYFPNFYALLK-----SISEKE-AKEKE 297
 Db 1386 QEOSLNKMKXDNETTKIVSEMEQFKPKDSALLRIEIMGLSKRLQESHDEMSVAKED 1445
 QY 298 TLTIIMKTLIDFVKMVMKYGTISPVEGVSYLENLDEMIA--LQTKNLE-----KNATD 349
 Db 1446 DLQRLQEVL-----QSSDQKLENIKKIVAKHLETEBELVAHCLKEQBE 1491
 QY 350 NISKLPAPSEKSHETDSTKKEAAKMEKYGSLKSDTKDSDNSNPGKTDPEKGTKE---406
 Db 1492 TINELRVNLSK-----ETEISTIQKLEAINDKLQN-----KIQEIVEKEEQLN 1536
 QY 407 -AYLEAIRKNIEWLKHDKGNKEDVDL-----SKMRDFINKQADAYVEKGLDKKEEAIAI 461
 Db 1537 IKQISEVQENVNELKQPKHRRKAKDSALQSIKSMLELTNRLQESQEIQIMIKEK-EEM 1595

Qy 462 KRIYSSL 468
||: :|
Db 1596 KRVQEAL 1602

Search completed: January 2, 2004, 11:00:01
Job time : 20 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 2, 2004, 10:59:27 ; Search time 40 Seconds
(without alignments)
3019.215 Million cell updates/sec

Title: ABU66614
Perfect score: 2399
Sequence: 1 MGFLGTGTTWLVLPQAF.....EKGILDKKEAEAKRIYSSL 468

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL 23.*
1: sp_archaea.*
2: sp_bacteria.*
3: sp_fungi.*
4: sp_human.*
5: sp_invertebrate.*
6: sp_mammal.*
7: sp_mhc.*
8: sp_organelle.*
9: sp_phage.*
10: sp_plant.*
11: sp_rodent.*
12: sp_virus.*
13: sp_vertebrate.*
14: sp_unclassified.*
15: sp_virus.*
16: sp_bacteriap.*
17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2062	86.0	466	11 Q8R1D7	Q8R1D7 mus musculus
2	1216	50.7	457	13 Q91847	Q91847 xenopus lae
3	198	8.3	7210	5 Q9V7G8	Q9V7G8 drosophila
4	198	8.3	9270	5 Q8MLD9	Q8MLD9 drosophila
5	197	8.2	2081	10 Q9LH98	Q9LH98 arabidopsis
6	196.5	8.2	1175	16 Q8XNW6	Q8XNW6 clostridium
7	196	8.2	2139	5 Q07569	Q07569 entamoeba h
8	191.5	8.0	2760	5 Q815Y2	Q815Y2 plasmodium
9	189	7.9	1434	5 Q8I492	Q8I492 plasmodium
10	185.5	7.9	1661	5 Q06166	Q06166 plasmodium
11	184	7.7	951	5 Q96229	Q96229 plasmodium
12	183.5	7.6	1790	3 Q07380	Q07380 saccharomyc
13	183.5	7.6	2612	5 Q815X5	Q815X5 plasmodium
14	180.5	7.5	3504	5 Q8IL45	Q8IL45 plasmodium
15	179.5	7.5	1320	11 Q9JK25	Q9JK25 rattus norv
16	179.5	7.5	1510	5 Q25920	Q25920 plasmodium

17	179	7.5	853	10 Q9LFE4	Q9LFE4 arabidopsis
18	179	7.5	1558	5 Q96275	Q96275 plasmodium
19	179	7.5	1786	5 Q9U0P0	Q9U0P0 plasmodium
20	179	7.5	2269	5 Q26223	Q26223 plasmodium
21	179	7.5	2747	5 Q9BJX9	Q9BJX9 plasmodium
22	178.5	7.4	1323	5 Q9NB35	Q9NB35 plasmodium
23	178.5	7.4	3317	16 Q8SWP8	Q8SWP8 mycoplasma
24	178	7.4	840	5 Q815X4	Q815X4 plasmodium
25	177.5	7.4	1464	5 Q81IF6	Q81IF6 plasmodium
26	177	7.4	1112	16 Q8EWQ1	Q8EWQ1 mycoplasma
27	177	7.4	1871	10 Q9SRD5	Q9SRD5 arabidopsis
28	176.5	7.4	682	5 Q8IBV6	Q8IBV6 plasmodium
29	175.5	7.3	1964	5 Q8SWQ7	Q8SWQ7 loligo peal
30	175	7.3	10578	5 Q8ISF5	Q8ISF5 caenorhabdi
31	175	7.3	18519	5 Q8ISF6	Q8ISF6 caenorhabdi
32	175	7.3	18534	5 Q8ISF7	Q8ISF7 caenorhabdi
33	174.5	7.3	924	5 Q15738	Q15738 dictyosteli
34	174.5	7.3	1263	5 Q8IK49	Q8IK49 plasmodium
35	173.5	7.2	1795	5 Q8IE35	Q8IE35 plasmodium
36	172.5	7.2	1134	2 Q99051	Q99051 streptococ
37	172.5	7.2	1979	5 Q96133	Q96133 plasmodium
38	172	7.2	495	5 Q8IEK6	Q8IEK6 plasmodium
39	171.5	7.1	2083	5 Q9N435	Q9N435 caenorhabdi
40	171.5	7.1	3484	5 P91257	P91257 caenorhabdi
41	171.5	7.1	5507	5 Q8IHN3	Q8IHN3 plasmodium
42	170.5	7.1	2771	5 Q26216	Q26216 plasmodium
43	170	7.1	1140	4 Q9ULE5	Q9ULE5 homo sapien
44	170	7.1	1804	11 Q921Q1	Q921Q1 mesocricetu
45	170	7.1	3130	5 Q9BK46	Q9BK46 plasmodium

ALIGNMENTS

RESULT 1

Q8R1D7 PRELIMINARY; PRT; 466 AA.
AC Q8R1D7;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE Similar to secretogranin III.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OC NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC024785; AAH24785.1; -;
SQ SEQUENCE 466 AA; 52783 MW; AB1CB9667276315C CRC64;

Query Match 86.0%; Score 2062; DB 11; Length 466;
Best Local Similarity 87.0%; Pred. No. 2.1e-94;
Matches 410; Conservative 17; Mismatches 36; Indels 8; Gaps 2;

Qy	1	MGFLGTGTTWLVLPQAFKPGSGQDSKLNHRELSEAPLNQIAEAEADKIKKT	57
Db	1	MGFLWTGTTWLVLPVLSNGPIQAFKPGSQDSKLNHRELSEAPLNQIAEAEADKIKKA	60
Qy	58	YPENKPGQSNYSFVDNLNLKAITKEKEKESQSRSSPLDNKLNVEDVDSTGNKLI	117
Db	61	FPESKPSSESNYSVDNLNLRLAITKETVEKERQSRSPFDNQLNVEDDSTGNKLI	120
Qy	118	DDYDSTKSGLDHFKQDDPDLGDLHLDGTPPLPAEDIVHKIAARIYEENDRAVDFDKIVSKLN	177
Db	121	DEYDSTKSGLDHFKQDDPDLGDLHLDGTPPLPAEDIVHKIATRIYEENDRGVDFDKIVSKLN	180
Qy	178	LGILITESQAHITLDEVAEVLQKLIISKANNYEEDPNKPTSWTENQAGKIKEKVTPTMAAIQ	237
Db	181	LGILITESQAHITLDEVAEALQKLIISKANNYEETLDKPTSRTEENQDGKIKEKVTPTVAAVQ	240

```
Db 238 DGLAKGENDVSNLTLTNGLERTKTYSNDFELOYFPNFFYALLKSIDSEKAKE 297
Qy 241 DGFTNRENDVSNLTLTNGLERTRNPHREDDFELQYFPNFFYALLTSIDSEKAKE 300
Db 298 TLITIMTKLIDFVQWVKYGTISPPEGVSYLENLDEMIALQTKNLEKNATDNISKLP 357
Qy 301 TLITIMTKLIDFVQWVKYGTISPPEGVSYLENLDEMIALQTKNLEKNATDNISKLP 355
Db 358 PSEKSHETDSTKEAAKWEKYSGLKSTKDDNSNPGKTDPEPKTEAYLEAIRKNIE 417
Qy 356 PPEKSEETDSTKEAAKWEKYSGLKSTKDDNSNPGKTDPEPKTEAYLEAIRKNIE 415
Db 418 WLKXKDKGKNEDYDLKQWDFINQADAYVEKGLDKKEEAIAKRIYSSL 468
Qy 416 WLKXKDKGKNEDYDLKQWDFINQADAYVEKGLDKKEEAIAKRIYSSL 466
Db 416 WLKXKDKGKNEDYDLKQWDFINQADAYVEKGLDKKEEAIAKRIYSSL 466

RESULT 2
Q91847 ID Q91847 PRELIMINARY; PRT; 457 AA.
AC Q91847;
DT 01-NOV-1996 (TremBLrel. 01, Created)
DT 01-NOV-1996 (TremBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TremBLrel. 19, Last annotation update)
DE Secretogranin III.
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
OC Xenopodinae; Xenopus.
OX NCBI_TaxID=8355;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Int. pituitary;
RX MEDLINE=96217355; PubMed=8632145;
RA Holthuis J.C.M., Martens G.J.M.;
RT "The neuroendocrine proteins secretogranin II and III are regionally
RT conserved and coordinately expressed with proopimelanocortin in
RT xenopus intermediate pituitary.";
RL J. Neurochem. 66:2248-2256(1996).
DR EMBL; X92872; CA63478.1; -
SQ SEQUENCE 457 AA; 51903 MW; B3096512D4B076D7 CRC64;

Query Match 50.7%; Score 1216; DB 13; Length 457;
Best Local Similarity 57.0%; Pred. No. 1.2e-52;
Matches 269; Conservative 63; Mismatches 106; Indels 34; Gaps 10;

Qy 8 TWILVLPIQAPKPGSQKSLHNRSLAERPLNEQIAEAEKIKKTYPPENKPGQS 67
Db 9 TAIIGVFHVQGFNPSPKQDKVHNRSLSEERPLEEQIAEA--DTVSRGTSTENQMLK 66
Qy 68 NYSFVDNLNLKATEKEKIEKERSIRSSPLDNKLVNEDVDSTKVRKLIIDYDSTKSL 127
Db 67 NDSFADDLGLFWSGKEDGDVRSSTGSGFSGNRGLTDDADSTKVLKLABYFDTTSM 126
Qy 128 DHKQDDPDGLHQLDGTPLTAEIVHKIAARIYENDRAVDKIVSKLNLGLITESQAH 187
Db 127 DYKEDDPDGLHQLD-APTAEIVRKTIARIYENDRGVDFKIVSKLNLGLITESQAY 185
Qy 188 TLEDEAVELQKLISKEANNVEE-----DPNKPTSWTENQAGKPEKVTPTMAAIO 239
Db 186 TLEDEAVELQQLIANEAKNEKEAGLDYSDVRSDDVDDDKREGKM-----ET 234
Qy 240 LAKGENDT--VSNLTLTNGLERTKTYSNDFELOYFPNFFYALLKSIDSEKAKE 297
Db 235 LNKNEDESSSTKNEDSLST-VERNELSPED---DLQYFPNFRVLKSLNENDVKEKK 290
Qy 298 TLITIMTKLIDFVQWVKYGTISPPEGVSYLENLDEMIALQTKNLEKNATDNISKLP 357
Db 291 TLLTIMTKLIDFVQWVKYGTITPEEGVNYLENLDDMIATVQVKNLGNFGSHRNAL--- 347
Qy 358 PSE-KSHETDSTKEAAKWEKYSGLKSTKDDNSNPGKTDPEPKTEAYLEAIRKN 416
Db 356 PPEKSEETDSTKEAAKWEKYSGLKSTKDDNSNPGKTDPEPKTEAYLEAIRKN 415
```

```
Db 348 PSDEKNADESDSTKEAAKWEKYSGLKSTKPEPND--AAESPKGKAETYLEAIRKNI 405
Qy 417 EWLKXKDKGKNEDYDLKQWDFINQADAYVEKGLDKKEEAIAKRIYSSL 468
Db 406 EWLKXKDKGKNEDYDLKQWDFINQADAYVEKGLDKKEEAIAKRIYSSL 457

RESULT 3
Q9V7G8 ID Q9V7G8 PRELIMINARY; PRT; 7210 AA.
AC Q9V7G8;
DT 01-MAY-2000 (TremBLrel. 13, Created)
DT 01-OCT-2002 (TremBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TremBLrel. 23, Last annotation update)
DE CG18255 protein.
GN STRN-MLCK OR CG8304 OR CG18255.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkely;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Banton R.C., Rogers Y.-H.C., Blazew R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Botchan M.R., Boudreau N., Boudreau P., Brotherton P.,
RA Borkova D., Botchan M.R., Boudreau N., Boudreau P., Brotherton P.,
RA Burtis K.C., Buser D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Fodor C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Gilek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laoko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Matti L., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacle J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu S.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,
RA Banton R.C., Baxton J., Baxton J., Beeson K.Y., Buser D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferreira S., Frase E., Galle R.F., Garg N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
```

[illegible]

DB 5127 DENDKK-QKEDGATNKSQK-AEADVVPEK-15EBEKVAEIK 5164

RESULT 4

Q8MLD9

ID Q8MLD9 PRELIMINARY; PRT; 9270 AA.

AC Q8MLD9;

DT 01-OCT-2002 (TrEMBLrel. 22, Created)

DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)

DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)

DE CG18255-PA.

GN STRN-MLCK OR CG8304 OR CG18255.

OS Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

ON NCBI_TaxID=7227;

OX [1]

RP SEQUENCE FROM N.A.

RC STRAIN=Berkeley;

RC MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,

RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,

RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,

RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,

RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,

RA Fostier C., Gabriellian A.S., Garg N.S., Gelbart W.M., Glaeser K.,

RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,

RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,

RA Nelson D.R., Nelson K.A., Nixon K., Nusekern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J.R., Puri V., Resse M.G.,

RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,

RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissensbach J.,

RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,

RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;

RT "The genome sequence of *Drosophila melanogaster*."

RL Science 287:2185-2195(2000).

RL [2]

RLN SEQUENCE FROM N.A.

RP Celniker S.E., Adams M.D., Kronmiller B., Wan K.H., Holt R.A.,

RA Evans C.A., Gocayne J.D., Amanatides P.G., Brandon R.C., Rogers Y.,

RA Banzon J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,

RA Carlson J.W., Center A., Champagne M., Davenport L.B., Dietz S.M.,

RA Dodson K., Dorsett E., Doup L.E., Doyle C., Dresnek D., Farfan D.,

RA Ferrara S., Frise E., Galle R.F., Garg N.S., George R.A.,

RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,

RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,

RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,

RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,

RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,


```

Db 1009 --NREKKYEEKSKTKBEAKKKSQDKKREKDSERKSKKEBSRDLKAKKKEE 1066
QY 401 PKGKTEAYLEAIRKNIEWLKHKDKGKEDYDLSDRDFINKQADAVYVEKGILDKKE 457
Db 1067 TREKKES-----ENHKKKKKKKHEHDNKSMMKKEE--KKEKKKHESKSRKKEE 1115

RESULT 6
QBXNW6 PRELIMINARY; PRT; 1175 AA.
AC Q8XNW6;
DT 01-MAR-2002 (TReMBLrel. 20, Created)
DT 01-MAR-2002 (TReMBLrel. 20, Last sequence update)
DT 01-MAR-2003 (TReMBLrel. 23, Last annotation update)
DE Probable exonuclease.
GN SBCC OR CPE0216.
OS Clostridium perfringens.
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;
OC Clostridium.
OX NCBI_taxid=1502;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=13 / Type A;
RX PubMed=11792842;
RA Shimizu T., Ohtani K., Hirakawa H., Ohshima K., Yamashita A.,
RT Shiba T., Ogasawara N., Hattori M., Kuhara S., Hayashi H.;
RT "Complete genome sequence of Clostridium perfringens, an anaerobic
RT flesh-eater.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:996-1001(2002).
DR EMBL; AP003185; BAB79922.1; -.
DR InterPro; IPR003439; ABC_transporter.
KW Exonuclease; Complete proteome.
SQ SEQUENCE 1175 AA; 136878 MW; 7C6D2366525019C1 CRC64;

Query Match 8.2%; Score 196.5; DB 16; Length 1175;
Best Local Similarity 22.6%; Pred. No. 0.078;
Matches 116; Conservative 83; Mismatches 196; Indels 119; Gaps 21;

QY 29 KSLHNRLSAERPLNEQIAEAEEDKIKTYPENKPGQSNYSFVDNL-NLLKAITEKE-- 85
Db 256 KELYDKRIEEB-----SLVSRSEIEKSPERVEISNKADKIVFINNLEILKEINKEDLK 311
QY 86 --KIEKQSRSSPLDNKLNVDVSTKNRKLIDDYDSTKSGLDHKFDQDPDGLHLDG 143
Db 312 FSELNKLLEELINLEENKLFKEPTKKEKLPDLRLKKEKLLS--QKRDILFOIKA 369
QY 144 TPLTAEDIVHKI-----AARIYEENDRAVDFKIVSK-----LNLG 179
Db 370 DGVKLKEACKLIFEDRSKCDTKLNSIEENKRLNEELKEEKEERKEELFVHEFNKINS 429
QY 180 LITESQAHTLEDEVAEVLQKLI--SKEANNYEEDPNKPTSWTENQAGKIPKVTPMMAIQ 237
Db 430 LFILNSYSLDKQFNEIKSEEVELKKYIKNLTEDEKSEK-----DLKVKVSLSKIR 482
QY 238 DGLAK-----GENDETVSNTLTNLGLERRTKTYSNDFEE-LQYF 277
Db 483 DKLESLLKETPGDSNSILEKQIKLGEYREKLNKYKEIKNSLEESLKT--KNFBEKLT 540
QY 278 PNFYALLASIDSEKAEKETLITIMKT--LIDFVKMVKYGTISPPEGVSYLE----- 329
Db 541 ENQKILL-----EKEVRELKDYINKVVEELAHKLRENLEVEGECPCVCGSTHHELNK 595
QY 330 -NLDE---MIALQTKNLEKNATDNISKL---FPAPSEK-----SHEETDSTKEAAK 375
Db 596 INLESNEKTLILESKEKELKELILEFSKIEATLEYENKKLEBELNISIEEVEGVEEERLK 655
QY 376 -MEKEYGSLDKSTKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLKHKDKGKEDYDL 434
Db 656 FLEBEFNTLKDIEEFNLK-----ENLEKLEKLLK--EENKNLE----- 693
QY 435 KMRDFINKQADAVYVEKGILDKKEEAPAIKRIYSSL 468
Db 694 ---NIFNKAERVILCEKIVREKEIASIKELDKEL 724

```

```

RESULT 7
Q07569 PRELIMINARY; PRT; 2139 AA.
AC Q07569; Q02504;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-OCT-2002 (TReMBLrel. 22, Last annotation update)
DE Myosin heavy chain.
GN MHCA.
OS Entamoeba histolytica.
OC Eukaryota; Entamoebidae; Entamoeba.
OX NCBI_taxid=5759;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=HMI:IMSS;
RX MEDLINE=93295430; PubMed=8515774;
RA Raymond-Denise A., Sansonetti P., Guillen N.;
RT "Identification and characterization of a myosin heavy chain gene
RT (mhca) from the human parasitic pathogen Entamoeba histolytica.";
RL Mol. Biochem. Parasitol. 59:123-131(1993).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=HMI:IMSS;
RA Guillen N.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
DR EMBL; L03534; AAB48065.1; -.
DR HSSP; P08799; IMND.
DR InterPro; IPR000048; IQ_region.
DR InterPro; IPR001609; myosin_head.
DR Pfam; PF00612; IQ; 2.
DR PRINTS; PR00193; MYOSINHEAVY.
DR ProDom; PD000355; myosin_head; 1.
DR SMART; SM00242; MYSC; 1.
DR SEQUENCE 2139 AA; 245225 MW; C68307341DBS1DD1 CRC64;

Query Match 8.2%; Score 196; DB 5; Length 2139;
Best Local Similarity 21.0%; Pred. No. 0.17;
Matches 107; Conservative 108; Mismatches 191; Indels 104; Gaps 23;

QY 20 FPKPGSQDKSLHNRLSAERPLNEQIAEAEEDKIKTYPEN-----KPGQSNYSFV 72
Db 859 FEEGKKKKKEIED-----LKKLAE---EIKKREAAENALASATKTGLEAKIQ 906
QY 73 D-----NLLKAITEKEKEKQSRSSPLDNKLNVDV-- 109
Db 907 DLEDKISELSKLSAAELDKQELNKLIENTLEEDKEELKTDIDNLKDKSKLKGDELEV 966
QY 110 -----STKNRKLDDVDSTKSGLDHKFDQDPDGLHLDGTPLTAEEDIVHKIAARIY 160
Db 967 EITELNSQIWTLN-ATVDKKDKTTIAEQESIDEKEDKIKLKGDIKLEEKDDL----- 1020
QY 161 EENDR-----AVFDKIVSKLNLGLITESQAHTLEDEVAEVLQKLISKEANNYEEDPNKPT 216
Db 1021 -EQDRAVDSATKDDIAKLNKTI-----ECEDAKDEIAKLEQEELEDEE--NKNKDLTNEL 1073
QY 217 SWTENQAGKIPKVTPMMAI-----ODGLAKG-ENDETVSNTLTNLGL-LEERTKTY 267
Db 1074 QQTOLKLGTEKSLAAQVAATKASDERDTLSQLENEKLTITKLTAKDLKLEKISGLK 1133
QY 268 EDNFEELQYFPNFYALLKSID-----SEKAEKETLITIMKTLDIFV-KMMVKYGTISP 321
Db 1134 QD-YEDDEDDKN-----KIEGDLRNAQRIKEUDDITGADVSQLQKQKEEY----- 1181
QY 322 EEGVSYLENLEDMIALQTKNLEKNATDNISKLPAPSEKSSHEETDSTKEAAK-----ME 377
Db 1182 ESQIAKMQEKEKAGNDVKNK-EKTIKE--KELEIQSLQKLEDETEVEKEDEAKKKEIE 1238
QY 378 KEYGSLDKSTKDDNSNPGGKTDEPKGTEAYLEAIRKNIEWLKHKDKGKEDYDL-SKM 436
Db 1239 KEMKALQEE-KENVESSKNSTKDKKKLEEDNLKDTQKLDMDTADNEKLEKAKAKDLQAOL 1297

```

QY 437 RDPFINKQADAYVEKGILDKKEAEAKRIYS 466
 Db 1298 NEVDNHEKAVADAELNKKQASDKELNS 1327

RESULT 8
 Q815Y2
 ID Q815Y2 PRELIMINARY; PRT; 2760 AA.
 AC Q815Y2;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Hypothetical protein.
 GN PFL0315C.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1] _SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RC STRAIN=3D7;
 RX MEDLINE=22255705; PubMed=12368864;
 RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W.,
 RA Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K.,
 RA Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S.,
 RA Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S.,
 RA Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B.,
 RA Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A.,
 RA McFadden G.I., Cummings L.M., Subramanian L.M., Mungall C.,
 RA Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W.,
 RA Fraser C.M., Barrrell B.;
 RT "Genome sequence of the human malaria parasite Plasmodium
 RT falciparum."
 RL Nature 419:498-511(2002).
 DR EMBL; AE014845; AAN36152.1; --
 KW Hypothetical protein.
 SQ SEQUENCE 2760 AA; 330896 MW; EE9964C4845181AC CRC64;

Query Match 8.0%; Score 191.5; DB 5; Length 2760;
 Best Local Similarity 21.1%; Pred. No. 0.37;
 Matches 121; Conservative 92; Mismatches 207; Indels 153; Gaps 24;

QY 29 KSLHNRELSAERPLNEQIAEAE-----DKIKTYPENKPGQSNVGFVD 73
 Db 1452 KNEAKEEETEKNQDNQVNEKEMDVNSKNREIVQVHNEIKNTNNKEEGKKNLLKEK 1511

QY 74 NIN--LLKAITEKIKERQ--SIRSPDLNKLNVDSVSTKNRLIDYDSTKGLDH 129
 Db 1512 EINDCLNDYINKQKKEKKNNWAMYGRPIVKRQNNRNINIKNDLKLKLYSSKESGFNDY 1571

QY 130 KF-----QDDPDGLHQLDGTPLTAE--DIVH-----KTAARIYEEN-- 163
 Db 1572 AFYAERFEVITGYNSEFDYSLDIDNOAKNEKNNDIHHNNIIKSKWKENIYENSPF 1631

QY 164 ---DRAVED-----KIVSKLNLGLITESQA----- 186
 Db 1632 HTYGRPIYEKSKNPNYNNKISTHNAILKKRKTLLAKSISFTQNSNNKIVR 1691

QY 187 -----HTLEDEAVLVOKLISKANNYEEDP-----NKPTSWTENQAGK 225
 Db 1692 TSIGNNTIDYNNSTIKIHKQN--VEDQGYIDLTKRKLTYDALDEINQOQKNLK 1749

QY 226 IPEKVTPTMAIQLAGENDVTSNLTITNG--LERRTKYSENPEELQYFNFYAL 283
 Db 1750 ITENIT-----QVQKHGQN---VSNIIKNTGAMLEKTRK-GNDNNEDQFSEELKAL 1800

QY 284 LKSIDSEKAKEKETITIMKTLIDFVQWVKYGTISPEEGVSYLENDEM-----I 335
 Db 1801 EK-LKKLKELKTEELKTEE--EKKRLLELNSLKVEEKKKKKKKMEEMKMEEMKI 1856

QY 336 ALQTKNLEK-----NATDNI SKL-----FPAPSEKSHBEETSTKEEAAKM--EKEK 380
 Db 1857 ELQKKEEELKQYRRQOIRKLELKKKEELKLEEMNKLKEEQRKKEIKKMEEKER 1916

QY 381 GSKLDSTKODNSNPGGKTDEPKGKTAYL-----FAIRKNIS--WLKHKDKKGNKEDY 431
 Db 1917 EKLKIKKEQKQKEEMKKQEQKKEQKLMMEEMRKFQEKVKVLEBEEKKEELKKL 1976

QY 432 DLSKMRDFINKQADAYVEKGILDKKEAEAKRI 464
 Db 1977 EQKKRE-ERKQWEEBEEKKEEKKKEEIKKM 2008

RESULT 9
 Q81492
 ID Q81492 PRELIMINARY; PRT; 1434 AA.
 AC Q81492;
 DT 01-MAR-2003 (TrEMBLrel. 23, Created)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
 DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
 DE Mature parasite-infected erythrocyte surface antigen (MESA) or
 DE PfEMP2.
 GN MESA.
 OS Plasmodium falciparum (isolate 3D7).
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxID=36329;
 RN [1] _SEQUENCE FROM N.A.
 RP SEQUENCE FROM N.A.
 RA Devlin K., Baker S., Davies P., Mungal K., Berriman M., Pain A.,
 RA Hall N., Bowman S., Churcher C., Quail M., Barrrell B.;
 RL Submitted (SEP-2002) to the EMBL/GenBank/DBSJ databases.
 [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22255708; PubMed=12368867;
 RA Hall N., Pain A., Berriman M., Churcher C., Harris B., Harris D.,
 RA Mungall K., Bowman S., Atkin R., Baker S., Barron A., Brooks K.,
 RA Buckee C.O., Burrows C., Cherevach I., Chillingworth C.,
 RA Chillingworth T., Christodoulou Z., Clark L., Clark R., Corton C.,
 RA Cronin A., Davies R., Davis P., Dear P., Dearden F., Doggett J.,
 RA Fellwell T., Gobie A., Goodhead I., Gwilliam R., Hamlin N., Hance Z.,
 RA Harper D., Hauser H., Hornsby T., Holroyd S., Horrocks P.,
 RA Humphray S., Jagels K., James K.D., Johnson D., Kerhornou A.,
 RA Knights A., Konfortov B., Kyes S., Larke N., Lawson D., Lennard N.,
 RA Line A., Maddison M., McLean J., Mooney P., Moule S., Murphy L.,
 RA Oliver K., Ormond D., Price C., Quail M.A., Rabinowitsch E.,
 RA Rajandream M.A., Rutter S., Rutherford K.M., Sanders M., Simmonds M.,
 RA Seeger K., Sharp S., Smith R., Squares R., Squares S., Stevens K.,
 RA Taylor K., Tivey A., Unwin L., Whitehead S., Woodward J.,
 RA Sultston J.E., Craig A., Newbold C., Barrrell B.G.;
 RT "Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13."
 RL Nature 419:527-531(2002).
 DR EMBL; AL929351; CAD51374.1; --
 SQ SEQUENCE 1434 AA; 168287 MW; AB0005F9DC26C989 CRC64;

Query Match 7.9%; Score 189; DB 5; Length 1434;
 Best Local Similarity 21.1%; Pred. No. 0.23; Mismatches 101; Conservative 96; Indels 102; Gaps 23;

QY 27 QDKSLHNRELSAERPLNEQ-----IAEAEEDKIKKTYPPENKPGQSNVFDNLNLKAI 81
 Db 501 KDKVLGEGDKEDVKENDEQKDKVLGEGDKEDVK-----EKNDGKDK-KVIGSEKTOKEI 554

QY 82 TEKEKIEKE-----RQISRSPLDNKLNVDSVSTKNRLIDYDSTKGLDHKFPD--- 133
 Db 555 --KEKVEKRVKKCKKKVKKGIKEN--DTEGNDKVGKEPII--IEEVKEEIKKQVEDGIK 608

QY 134 --DPDGLHQLDGTPLTAEIVHKAARIYE--ENDRAVDFKIVSKLNLGLITESAHT 188
 Db 609 ENDETEGNDKVGKEPIITEEVKEEIKKQVEGIEKENDTEGNDKVGKP-----EITEEVKEE 664

QY 189 LEDEVAEVLQ-----KLISKEANNYEEDPNKPTSWTENQAGKIPKVTPTMAAIQDGL 240
 Db 665 IKKQVEGIEKENDTESKDLIGQBIITEEVKEGIEKENDTENK-----DKVIGQBIITEEV 719

QY 241 AKG--ENDETVSNLTITNGLERRTKYSDNEFEEQLQYFNFYALLKSIDSEKAKEKET 298

Db 720 KEGIKEND-----TENKDKVIGQEIITE-----EYKKSIEQKEKNGKEN 759

QY 299 LITI-----MKTLDIFVMMVKYGTISPBGV--SYLENLD-----EMIALQTKNKLK 345

Db 760 ILEIKDIVIGQEVIIIEVKVKIKK--KVEKGIKENHTSEKDKVIGQEVIIIEVKKEIEK 816

QY 346 NATDNISKLFPAPSEKSEHEETSTKEEAAKKEVGSLSKSTKDDNSPGGKTDEPKGT 405

Db 817 QVEGIG-----ENDTESKDKVIGQEVIGK--DVNEEGPENKDKVTKQEKVK- 861

QY 406 EAYLEAIRKNIEMWLKHKDKGNKEDYDLSK--MRDFINKQADAYVEKGIIDKEEAEAIK 462

Db 862 EYKVEKVKVKVKKVKKRNKNKRNKDNVIGKELMKEDVNEKDTANKDKVIEQKEKEEYK 920

RESULT 10

Q06166

ID Q06166 PRELIMINARY; PRT; 1661 AA.

AC Q06166

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)

DT 01-OCT-2002 (TREMBLrel. 22, Last annotation update)

DE Mature PARASITE-infected erythrocyte surface antigen (Antigenic protein PREMP2).

DE Plasmodium falciparum.

OS Plasmodium falciparum.

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=5833;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=PALO ALTO;

RC MEDLINE=92158014; PubMed=1741020;

RA Coppel R.L.;

RT "Repeat structures in a Plasmodium falciparum protein (MESA) that binds human erythrocyte protein 4.1.";

RL Mol. Biochem. Parasitol. 50:335-347(1992).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=PALO ALTO;

RA Kun J.F.J., Waller K.L., Coppel R.L.;

RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE OF 797-850 FROM N.A.

RP MEDLINE=93122844; PubMed=1478701;

RA Saul A., Yeganeh F., Howard R.J.;

RT "Conservation of repeating structures in the PfEMP2/MESA protein of Plasmodium falciparum.";

RL Immunol. Cell Biol. 70:353-355(1992).

DR EMBL; AF056936; AAC33303.1; -.

DR InterPro; IPR001623; DnaJ_N.

DR Pfam; PF00226; DnaJ_1.

DR SMART; SM00271; DnaJ_1.

DR PROSITE; PS00076; DnaJ_2; 1.

KW Antigen.

SQ SEQUENCE 1661 AA; 195479 MW; AF340527D85A9D29 CRC64;

Query Match 7.9%; Score 188.5; DB 5; Length 1661;

Best Local Similarity 19.9%; Pred. No. 0.29;

Matches 94; Conservative 90; Mismatches 189; Indels 99; Gaps 15;

QY 22 KPGSQDKSLHNRLESAERPLNEQIAEAEEDKIKTKYPPENKPGQSNYSFVNLNLLKAI 81

Db 895 KDTANKKEIEQKEKEEVEKKEEVEKKEEVEKKEE-----V 937

QY 82 TEKEKIERQSRSSPLDNKLVNEDVDSTKNRKLIDDDYDSTKSLDHLKFDQDPGLHQL 141

Db 938 KEKEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEE 996

QY 142 DGTPLTADIVHKIAARYEEN-----DRAVFDKIVSKLNLGLITESQAHTLEDEVAE 195

Db 997 IGOEIIIEIKKEVKVKVKKRNKNKNKNKNKNKNKNKNKNKNKNKNKNKNKNKNKNKN 1056

QY 196 VLQKLISKEANNYEEDPNKPTSWTENQAGKIPEKVTPTMAAIQDGLAKGENDETNTLTL 255

Db 1057 VKEKEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEE 1081

QY 256 TNGLERTKTYSEDNFEELQVFPNFYALLKSIDSEKAKEKET-----LITIMKTLDIPVK 311

Db 1082 -----KEKDTESKDKVIEQ-----KEKEVEKKEKDTENKDKVIGQEVIIIEIK 1127

QY 312 MMVKYGTISPBGVSYLENDEMIALQTKN-LEKNATDNISKLFPAPSEKSH-EETDST 369

Db 1128 KEVKRV---KRRNKN 1184

QY 370 KEEAAKKEVGSLSKDS-----TKDDNSNPGGKTDEPKGT---EAYLEATKRN 415

Db 1185 KEKEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEEVEKKEE 1244

QY 416 IE--WLKHKDKGNKEDYDLSK--MRDFINKQADAYVEKGIIDKEEAEAIKRI 464

Db 1245 VKKEVKRRNKN 1296

RESULT 11

O96229

ID O96229 PRELIMINARY; PRT; 951 AA.

AC O96229

DT 01-MAY-1999 (TREMBLrel. 10, Created)

DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)

DT 01-MAR-2003 (TREMBLrel. 23, Last annotation update)

DE Hypothetical protein.

GN PFB0680W.

OS Plasmodium falciparum (isolate 3D7).

OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

OX NCBI_TaxID=36329;

RN [1]

RN SEQUENCE FROM N.A.

RC STRAIN=3D7;

RC MEDLINE=99021743; PubMed=9804551;

RA Gardner M.J., Tettelin H., Carucci D.J., Cummings L.M., Aravind L., Koonin E.V., Shallow S., Mason T., Yu K., Fujii C., Pederson J., Shen K., Jing J., Aston C., Lai Z., Schwartz D.C., Pertea M., Salzberg S., Zhou L., Sutton G.G., Clayton R., White O., Smith H.O., Fraser C.M., Adams M.D., Venter J.C., Hoffman S.L.;

RA "Chromosome 2 sequence of the human malaria parasite Plasmodium falciparum.";

RL Science 282:1126-1132(1998).

RN [2]

RN SEQUENCE FROM N.A.

RC STRAIN=3D7;

RC MEDLINE=2255705; PubMed=12368864;

RA Gardner M.J., Hall N., Fung E., White O., Berriman M., Hyman R.W., Carlton J.M., Pain A., Nelson K.E., Bowman S., Paulsen I.T., James K., Eisen J.A., Rutherford K., Salzberg S.L., Craig A., Kyes S., Chan M.-S., Nene V., Shallow S.J., Suh B., Peterson J., Angiuoli S., Pertea M., Allen J., Selengut J., Haft D., Mather M.W., Vaidya A.B., Martin D.M.A., Fairlamb A.H., Fraunholz M.J., Roos D.S., Ralph S.A., McFadden G.I., Cummings L.M., Subramanian G.M., Mungall C., Venter J.C., Carucci D.J., Hoffman S.L., Newbold C., Davis R.W., Fraser C.M., Barrell B.;

RA "Genome sequence of the human malaria parasite Plasmodium falciparum.";

RT Nature 419:498-511(2002).

RL EMBL; AB001410; AAC71925.2; -.

KW Hypothetical protein.

SQ SEQUENCE 951 AA; 112486 MW; AC8D889358A84F4F CRC64;

Query Match 7.7%; Score 184; DB 5; Length 951;

Best Local Similarity 22.0%; Pred. No. 0.25;

Matches 104; Conservative 76; Mismatches 182; Indels 110; Gaps 17;

QY 25 GSQDKSLHNRLESAERPLNEQIAEAEEDKIKTKYPPENKPGQSNYSFVNLNLLKAI 84

Db 161 GKQDINSNAE--NKKDVKEGVKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEEKEE 206

QY 85 EKIEKERQSRSSPLDNKLVNEDVDSTKNRKLIDDDYDSTKSLDHLKFDQDPGLHQL 140

[illegible]

```

RESULT 15
ID Q9JK25 PRELIMINARY; PRT; 1320 AA.
Q9JK25;
AC Q9JK25;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
CLIP-170.
DE CYLIN.
OS Rattus norvegicus (Rat.).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Hippocampus;
RA Galjart N.;
RL Submitted (MAR-1999) to the EMBL/GenBank/DBJ databases.
[2]
RN RN
RP SEQUENCE FROM N.A.
RC STRAIN=Wistar; TISSUE=Hippocampus;
RX MEDLINE=21185938; PubMed=11290329;
RA Akhmanova A., Hoogenraad C.C., Drabek K., Stepanova T., Dortland B.,
RA Verkerk T., Vermulen W., Burgering B.M., de Zeeuw C.I., Grosveld F.,
RA Galjar N.;
RT "CLASPs are CLIP-115 and -170 associating proteins involved in the
RT regional regulation of microtubule dynamics in motile fibroblasts.";
RL Cell 104:923-935 (2001).
DR EMBL; AJ237670; CAB92974.1; -.
DR InterPro; IPR000938; CAP-Gly.
DR InterPro; IPR001878; Znf_CCHC.
DR Pfam; PF01302; CAP Gly; 2.
DR SMART; SM00343; Znf_C2HC; 1.
DR PROSITE; PS00845; CAP GLY 1; 2.
DR PROSITE; PS50245; CAP_GLY 2; 2.
SQ SEQUENCE 1320 AA; 148264 MW; 87BE0F0463D20E9D CRC64;

Query Match 7.5%; Score 179.5; DB 11; Length 1320;
Best Local Similarity 22.6%; Pred. No. 0.62;
Matches 134; Conservative 77; Mismatches 211; Indels 171; Gaps 27;

QY 22 KPGSDOKSLH-NRELSAERPLNEQIAAEEDKIKKTYPPENKPGOSNYSFVDNLNLKA 80
Db 525 KPGDGVDMLSLLQLEISA---LQEKLVTHDQNEVTSLKQFGFGREEMFQKEIKALHA 581
QY 81 ITEKEKIEKEROSIRSSPLD--NKLNVEDV----- 108

```

